VRIJE UNIVERSITEIT AMSTERDAM

VU Research Portal

Economic evaluation of urban and spatial policies

Levkovich, O.D.

2017

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

citation for published version (APA)
Levkovich, O. D. (2017). *Economic evaluation of urban and spatial policies*. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam].

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl

Download date: 05. May. 2023

ECONOMIC EVALUATION OF URBAN AND SPATIAL POLICIES

ISBN:978-94-028-0826-1

Cover design by Rotem Fisch (*Rotem.fisch@gmail.com*)

VRIJE UNIVERSITEIT

ECONOMIC EVALUATION OF URBAN AND SPATIAL POLICIES

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de School of Business and Economics
op maandag 11 december 2017 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Or David Levkovich

geboren te Ramat-Gan, Israël

promotoren:	prof.dr.	J. Rouwendal
	prof.dr.	E.T. Verhoef

Contents

List of Tables

List of Figures

Preface

This PhD dissertation could not have been completed without the support of many colleagues, friends and family. Following a PhD trajectory was not something I imagined for myself five years ago. However, I was lucky to have inspiring teachers who encouraged me to discover interest in scientific research work and to rethink my plans.

I would first like to thank my primary promoter, Jan Rouwendal. It was with Jan's encouragement that I joined the department of Spatial Economics at the VU, and I owe much to his guidance. Our discussions and work together have taught me much on economic research, and it will certainly shape every new project that I will approach in the future. Thank you for seeing a potential in me, inspiring and guiding, while still giving me the freedom to try on my own (and often make necessary mistakes).

When I started my PhD my second promoter was Piet Rietveld. Piet passed away in November 2013, and I regret that I hardly had a chance to work under his guidance. His vast knowledge, broad interests and kind and humble character will remain a constant source of inspiration.

Erik Verhoef became my second promoter half-way through my PhD. Although we worked on relatively few projects together, I learned much from Erik's useful comments, attention to details and strive for excellence, but mostly from his availability to discuss and solve any issue.

I would also like to thank Jos van Ommeren. Discussing work together has always made me look at things from a different perspective, thank you for the useful comments and guidance along the way.

I am thankful to Elfie Bonke and Jenny Wiersema for being a great support for me in the past few years, and helping in almost everything from finding a place to live in Amsterdam to the preparation for graduation.

I would like to thank the other co-authors of the chapters of this dissertation - Ramona van Marwijk and Lars Brugman, as well as others with whom I had the pleasure of working with in the past years - Carla Sá, Edwin Buitelaar, Florian Sniekers and Ismir Mulalic.

The department of Spatial Economics has been a great place to grow pro-

fessionally. I thank the members of the department who contributed to the good working experience here, the pleasant and friendly atmosphere, and the interesting talks over coffee or beer during department borrels. I particularly like to thank Eric Koomen, Paul Koster, Hans Koster, Jasper Dekkers, Thomas de Graaff, Maureen Lankhuizen, Mark Lijesen, Eveline van Leeuwen, Niels van Manen, Hadewijch van Delft and Henri de Groot.

Working on the PhD I was also lucky to get to know many interesting people. I would like to particularly thank Jessie Bakens, Chris Behrens, Masha Dementyeva, Jan Möhlmann, Martijn Smit, Martin Adler, Alexandros Dimitropoulos, Francis Ostermeijer, Maurice de Kleijn, Zhiling Wang, Dani Broitman and Stefanie Peer for their help, and for making my past few years at the VU meaningful and enjoyable. I thank Mark van Duijn, my first officemate, for easing my arrival to the university. Ceren Ozgen, thank you for tolerating the noise in our office, constant supply of Turkish cookies, editing of scientific work, and your useful advices, friendship and support. Yuval Kantor, thank you for your valuable friendship. Your advices on academic work, cooking recipes, home improvement, parenting and more have made the whole experience much better.

Finally, I would like to thank my family. To my parents, Ron and Hana, my grandparents, Haviva and Jacob, and my brothers, Aner and Cnaan, thank you for the love, encouragement and constant support. I thank my son Jonathan, whose arrival put this work in perspective and gave it meaning. Lastly, to my wife, Noam, thank you for being a patient reader and a good listener. I am grateful for your love and support, without it this work could not have been completed.

Or Levkovich,
Amsterdam, August 2017

Chapter 1

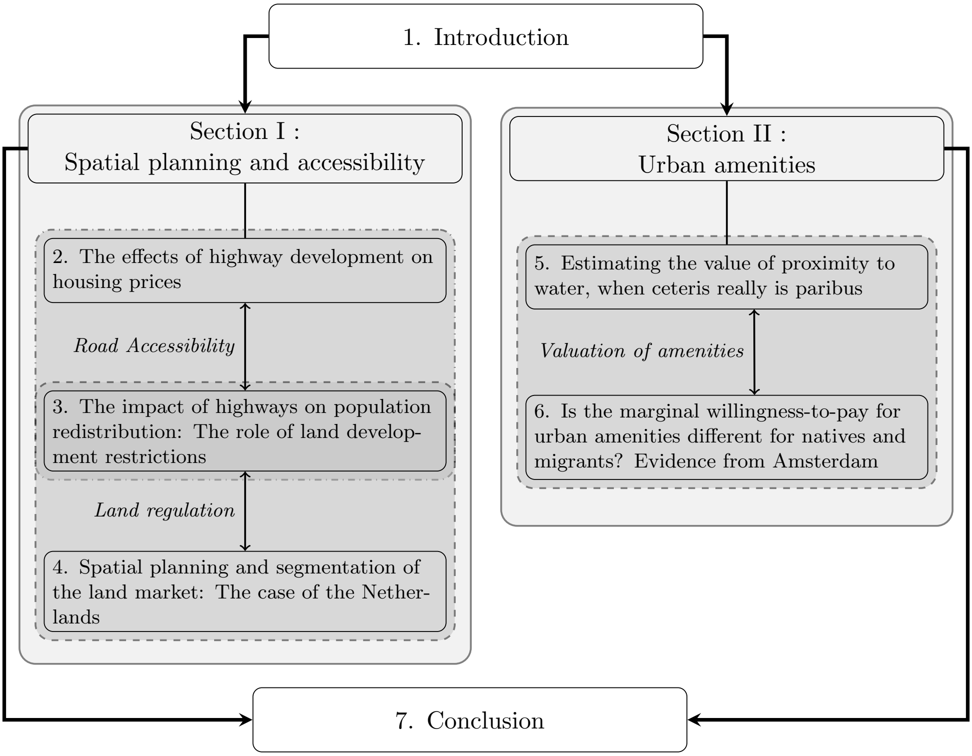
Introduction

This dissertation includes five research papers which use an empirical approach to explore the economic effects of urban spatial policies. Each of the chapters focuses on the impact of one or more urban spatial policies or decisions, among which are improved road accessibility, land development restrictions and land-use policies, and the provision of urban amenities. In the course of this dissertation I will demonstrate how the application of several urban policies in the Netherlands has had a substantial impact on housing and land markets, and on urban growth and the spatial distribution of households and individuals. Moreover, while most examined policies were effective in achieving their goals, they have also often resulted in unintended indirect outcomes, both in the short- and long-run.

The chapters of the dissertation are divided into two main sections, as illustrated in Figure 1.1. The first section focuses on the evaluation of the economic effects of spatial planning decisions and improvement in accessibility. The primary purpose of spatial planning systems is to guide and direct the growth and development of cities and regions. Such guidance of urban growth addresses current population needs and is done in order to improve social welfare. However, it also constitutes intervention and distortion in housing and land markets, which may consequently also affect the spatial distribution of population and economic activity.

It is often the case that such distortions are intentional or expected, as part of the policy's original purpose. However, such policies can often result in outcomes which are unexpected or contrary to the original policy's goal. For instance, while new transportation infrastructure can increase the demand for housing in newly accessible areas, such demand may remain unrealized if development restrictions are present. In the Netherlands, the de-concentration of the population into regional centers was facilitated by the expansion of the

Figure 1.1: Structure of the Dissertation



highway network and development restrictions in the vicinity of population centers. This consequence was later considered as an outdated policy, since it caused urban decline in traditional city centers and resulted in increased commuting distances and time (Schwanen et al., 2001; Geurs, 2006). An additional example is land development restrictions, which are placed in order to ensure orderly development of urban areas and preservation of open space, which is also valuable in itself (Vermeulen and Rouwendal, 2014). However, enforcement of land development restrictions can distort efficient allocation of land, increase local house prices and reduce welfare (Ihlanfeldt, 2007; Turner et al., 2014). Since such unintended outcomes may arise, understanding them is essential for evaluating the full welfare effect of those planning decisions.

Chapter two of the dissertation explores the effects of two highway development on house prices in rural areas in the Netherlands. To identify the causal effects of highway development on housing prices, I estimate a repeat-sales model with difference-in-differences estimator. Comparing transactions of the same properties before and after the construction of the new highways in both treated and controlled areas allows for better control for confounding factors and improved identification of this causal effect. The findings show that a new highway makes remote residential towns more attractive, and can significantly increase the value of homes. Moreover, the increase in real estate values begins already before the realization of the project, indicating that an anticipation effect is present.

The third chapter further examines the effects of new highways, under a setting in which strict land development restrictions are present. This is done in order to study how such land restrictions interfere with the effects of highways on suburbanization processes and the spatial distribution of population. In this chapter I demonstrate that spatial development restrictions can substantially interfere with these effects, and restrain urban development in the vicinity of new highways. To identify the effect of new highways on population growth we examine changes in population levels during the large-scale expansion of the Dutch highway network in the 1960's. We then follow an instrumental variable approach to address the endogeneity between roads expansion and population growth, we show that highways accelerated population growth in peripheral municipalities but did not affect population changes in central cities or suburbs. Moreover, the findings show that where urban development is restricted, new highways caused population to be distributed to areas further away from central location, compared with the predictions under an unregulated land market setting.

In the fourth chapter I examine the effects of land development restrictions, and try to answer how these restrictions, combined with strict land-use designations, affect the divergence in transaction prices of undeveloped land designated to different uses. This chapter involves an examination of land transactions data from the Netherlands, using fixed-effects analysis to identify the effect of land designation on the transaction value of undeveloped lands, as well as an analysis of municipal characteristics and agricultural land values in order to explain differences in this effect. The findings show that separate planning regimes for residential, commercial and agricultural land-uses are responsible for substantial price differences between undeveloped lands designated to these uses. Since conversion between uses is hardly possible, there exists a segmentation between the different land markets. This segmentation is reflected in a substantial divergence in prices of nearby land parcels designated to different uses.

The second section, which includes chapters five and six, focuses on the evaluation of policies which involve the provision of urban amenities. While it often relates to attractiveness, the provision of amenities defines the characteristics of the spatial setting, and is also closely related to the level of provision of necessary facilities for urban living (schools, for instance) as well as socio-demographic characteristic of a residential area. The level and quality of amenities, both natural and man-made, is shown to be an important determinant for location decisions of individuals (Glaeser et al., 2001; Brueckner et al., 1999), and it is reflected in various planning policies for new developments as well as in preservation of cultural heritage (Van Duijn and Rouwendal, 2013).

However, not all amenities are valued similarly, and heterogeneous individuals place different values for amenities. Since some amenities require larger investments (for instance, artificial water bodies or maintenance of cultural heritage), it is important to estimate whether they are indeed valued by local residents and workers. Knowledge of the value which individuals attach to amenities may improve the efficiency of such resource allocation.

In this section of the dissertation I address the provision and development of urban amenities as a form of urban policy, and estimate how they are reflected by the willingness-to-pay of households and individuals for amenities. These estimated values can be obtained from observed housing transaction prices and individual location decisions. The fifth chapter of the dissertation examines the value of proximity to water in new residential neighborhood in the Netherlands. The general consensus regarding the importance of amenities often also translates to official requirements for development projects of new residential neighborhood. In the Netherlands, large scale residential developments in the past decades were required to provide reasonable accessibility by public transport to employment centers, and to protect open space (Kruythoff and Teule, 1997; Boeijenga et al., 2008). In order to increase the attractiveness of new neighborhoods, planners also invested resources in developing additional amenities, such as artificial canals or lakes. However, if residents attach lower value than expected to such amenities it may be that the investment is inefficient. To address this we study proximity to water in a newly built neighborhood and use fixed effects for almost identical properties, which allows measuring the effect in almost perfect ceteris paribus conditions. The findings demonstrate that proximity to water has a positive effect on house prices only in immediate adjacency to water, and that this effect is substantially lower compared with previous findings[1].

The sixth chapter explores heterogeneity in preferences for urban amenities, focusing on highly-educated migrants to Amsterdam and their preferences for existing communities of migrants, historical features, nature, schools and accessibility. While urban amenities are important for the well-being of existing residents, they also play an important role in attracting new population (Bayer et al., 2004b, 2007; Klaiber and Phaneuf, 2010; Rodríguez-Pose and Ketterer, 2012). This role becomes particularly relevant for attracting highly-educated workers. Such workers are found to increase productivity and improve performance in skill-intensive industries, and to have a positive effect on regional employment and wages (Behrens et al., 2014; Glaeser and Saiz, 2003; Glaeser and Resseger, 2010). By analyzing observed location choices, and using residential sorting estimation, we investigate how the marginal willingness-to-pay patterns for urban amenities of highly-educated workers differ from those of locals and low- and medium-educated workers. The findings show that highly-educated migrants have stronger preferences for amenities, and are willing to pay more for higher provision of urban amenities compared with locals, all else equal.

The findings presented in this dissertation contribute to our understanding of the outcomes of urban spatial policies which are commonly applied in many regions worldwide. This is also particularly relevant for the design and implementation of future urban and spatial policies. A better understanding of the potential policy effects examined here can be used for several applications. These applications include quantifying the monetary value of policies, circumventing possible negative direct and indirect outcomes, avoiding misallocation of public resources and improving the effectiveness of policies.

[1] See for example Anderson and West (2006); Cho et al. (2006); Doss and Taff (1996).

Chapter 2

The effects of highway development on housing prices

2.1 Introduction[1]

Transportation infrastructure development is undertaken to improve accessibility at a regional or urban level and to relieve traffic congestion in these areas. This development is evaluated by property owners and residents in the affected areas, and is capitalized in the price of housing. However, the effect of new transportation development, roads or highways in particular, may have both positive and negative effects on the price of housing. Improved accessibility may shift housing prices upwards, whereas higher traffic noise levels and increase in traffic density may reduce prices in houses that are adjacent to the new road. Gaining an understanding of whether a development project is valued positively or negatively is a key issue in evaluating regional policy. Accurate estimations of such valuations may be used as a valuable criterion in

[1]This chapter is based on joint work with Jan Rouwendal (VU University Amsterdam and Tinbergen Institute, The Netherlands) and Ramona van Marwijk (Kadaster, the Dutch land registry). Apart from minor changes, this chapter is published as Levkovich, Or, Jan Rouwendal, and Ramona Marwijk. *"The effects of highway development on housing prices."* Transportation 43.2 (2016): 379-405.
I would like to thank Ronnie Lassche for his significant contribution in constructing the travel time database that formed the basis of the accessibility measure used in this chapter. I thank Kadaster, the Dutch land registry, for providing us with high quality data and insightful advice on study locations and development projects. I would also like to thank the participants of the Dutch–Israeli regional science workshop for a constructive discussion on the research.

a social cost-benefit analysis. Analyzing the effects of the various positive and negative externalities of the development project may also be used in project evaluation, and assessing which externality is most dominant and to what extent it affects housing prices has implications for future policy planning. The relevance of the research is highlighted by the disagreement among researchers concerning the total effects of transportation development on housing prices. Although researchers generally agree that positive and negative externalities exist at different proximities to a developed road, most findings differ to the extent that some reach opposite conclusions.

In this chapter, we apply a repeat sales and difference-in-differences method to investigate housing price dynamics following the development of two highways in the east of the Netherlands. We use a very large and unique housing transaction data from Kadaster, the Dutch Land Registry, which allows us to compare the development of housing transaction values over a long time-span and a large geographic area. The use of such data is essential for conducting a high-quality repeat-sales analysis, and it allows us to effectively control for housing and neighborhood characteristics which remain constant over time and to reduce concerns for omitted variable bias.

In the scope of the analysis, we focus on three main transportation development externalities: improved accessibility, increased traffic intensity and reduced noise levels. We argue that the effects of transportation infrastructure development on the price of housing vary between properties depending on their geographic location, municipality affiliation and proximity to the newly developed projects, and also on unobserved heterogeneity between properties. Our research focuses on several questions regarding the effect of transportation development on the price of housing. We first ask how infrastructure development is valued by owners and residents of the region's municipalities, and how they value the changes in the levels of each externality. In addressing this question, we estimate the elasticity of housing prices with respect to changes in accessibility, traffic intensity and noise pollution levels, and calculate the implicit willingness to pay (WTP) for changes in the levels of these externalities. The second question that is addressed in this research is how the strength and dominance of the effect of each of the transportation development externalities examined varies between different areas. As it is probable that transportation development affects housing prices differently in various geographic settings, such as in areas with differences in population densities, we control for differences between neighborhoods when estimating the effects of the highway development. Finally, considering all possible effects on housing prices, we estimate the total housing price effect in the affected municipalities and examine whether it has increased or decreased due to the construction of a new road. In addition, we use our estimation results to calculate WTP for accessibility improvements for each four-digit postal code in the study area.

The chapter is structured as follows. The next section contains a literature review. Section 2.3 presents the model specifications. Section 2.4 describes the data sources and data sets used in the research. Section 2.5 describes the results of the estimations and includes calculations of quasi-rents and WTP values for improved accessibility. Section 2.6 includes a conclusion and discussion regarding possible policy applications.

2.2 Literature review

The development of transportation infrastructure and the resulting drop in transportation costs and increase in accessibility levels are closely related to changes in housing values (Alonso, 1964). In the housing market literature, hedonic regressions are most commonly used to determine the price of a property as a function of its attributes (Rosen, 1974). The repeat sales method (Bailey et al., 1963) can be used as a modification of hedonic models, as well as to measure changes in a price of the same property sold over time. Originally developed as a method to construct price indices, the main advantage of the repeat sales method is that it allows fixed characteristics that influence property prices and were perhaps omitted from the hedonic regression to be ignored. Due to the wide application of both methods, different improvements have been introduced to address various challenges posed by the housing market and to reduce estimator bias (Wang and Zorn, 1997). One example is the use of the weighted repeat sales method, which was developed by Case and Shiller (1987); this approach considers changes in the variance of the error between observations and therefore yields more accurate estimations of housing price indices. To estimate the effect of external development or "treatment" on a group of houses, the difference-in-differences (DID) estimator may be incorporated as part of the repeat sales model (Imbens and Wooldridge, 2009). The DID estimator is used to compare the time effect between a group that was exposed to a treatment and a control group. The difference in time that is observed in the control group is then subtracted from the time of the treatment group to separate unrelated time appreciation effects from treatment effects. The inclusion of a DID estimator requires that the control and treatment groups are properly defined to reduce the likelihood of bias in the estimated coefficient (Imbens and Wooldridge, 2009).

The effects of transportation development on housing prices, in particular in relation to accessibility, have been addressed extensively in previous literature, predominantly using estimations of hedonic models (Armstrong and

Rodríguez, 2006; Cheshire and Sheppard, 1995; Coulson and Engle, 1987; Franklin and Waddell, 2003; Henneberry, 1998; Iacono and Levinson, 2012; Martínez and Viegas, 2009). Whereas most of these studies focus exclusively on the positive effects of improved accessibility on housing prices, other studies emphasize that both positive and negative externalities may result from highway development, and these may affect residential preferences (Debrezion et al., 2007; Iacono and Levinson, 2012; Martínez and Viegas, 2009; Tillema et al., 2012). Moreover, depending on the exposure levels to different externalities, house prices in different geographic areas may change differently due to the effects caused by the new transportation infrastructure development (Smersh and Smith, 2000). Negative effects may result from an increase in traffic noise pollution, which has also been found to be a cause of discount in the value of properties that are located along a newly developed highway (Kim et al., 2007; Nelson, 1982; Theebe, 2004; Wilhelmsson, 2000). Ossokina and Verweij (2015) study the effects of a new highway in The Hague on the surrounding residential properties using a repeat sales approach and focus particularly on the positive effects of the reduced traffic density. They find that property values in the proximity of the new road have increased with the reduction in traffic density, thus providing further evidence that homeowners value traffic density negatively. The effects of a new transportation development project have also been found to be reflected in housing prices even before its completion. Yiu and Wong (2005), Koster et al. (2010) and Cotteleer and Peerlings (2011) demonstrate that prices may adjust before project completion and that rational public expectations may account for the price capitalization of the new highway while it is still under construction.

The past literature raises several important issues. First, it highlights that a different focus on the positive or negative externalities of a transportation development can lead to differences in conclusions. Although researchers generally agree that positive and negative externalities exist at different proximities to the new or reconstructed road, most findings raise the suspicion of omitted variable bias as they do not consider the different effects of each externality over time and space. Other possible explanations for the difference in the dominance of externalities in the housing market may be the improper definition of control and treatment groups, or the spatial patterns of the sampled data. As demonstrated by Smersh and Smith (2000), the dominance of the effect of one externality over another depends greatly on the different spatial patterns in the sampled areas. Moreover, the past literature also shows that the design of the model requires special attention to avoid bias in the coefficients and to estimate properly the different effects on housing prices.

In this chapter, we address these challenges by estimating the effects of the development of two highways in the Netherlands on local housing prices using a repeat sales analysis combined with a DID approach. Aiming to reduce the suspicion of estimator bias, our analysis also considers that both positive and negative effects may be apparent in different municipalities in the environs of the highways. Furthermore, we take neighborhood effects into account and consider that prices may change prior to the completion of the highway development due to anticipation effects.

2.3 Model specifications

To determine the effects of transportation infrastructure development, we estimate three different versions of the repeat sales model. The three specifications correspond to the research questions and are used for comparison of the estimated coefficients and of model robustness. The repeat sales model specifications that we propose to use are: i) the traditional repeat sales model, including a DID estimator; ii) a specification that accounts for the three main transportation externalities and controls for neighborhood effects; iii) a specification that estimates the individual externality effects for each municipality. To control for spatial correlation in the residuals, the models are estimated using clustered standard errors, in which the clusters are defined based on six-digit postal codes. All suggested model forms derive from the traditional repeat sales form, introduced by Bailey et al. (1963):

$$\log \widetilde{P_i^{T,t}} = \sum_{t=0}^{T} \beta_i^t X_i^t + \epsilon_i^{T,t} \tag{2.1}$$

where $\widetilde{P_i^{T,t}} = (\frac{P_{iT}}{P_{it}})$ is the change in price of house i between times t and T. X_i^t is equal to -1 at the initial time of sale, +1 at the final time of sale, and 0 otherwise. As an addition to this model, we incorporate DID estimators as an interaction variable, which indicates that property i was sold during the time ($T_i = 1$) and is located in the treatment area group ($G_i = 1$): $I_i = T_i * G_i$. The model specifications also take anticipation effects into account. As the construction of both projects began in 2001, the anticipation period was set to begin in the year 2000 to capture the full effects of public expectation. It is likely that public anticipation in relation to the development of the projects began slightly before the start of construction, and therefore properties that were sold between the years 2000–2004 were included in the anticipation period. Thus:

$$\log \widetilde{P_i^{T,t}} = \sum_{t=0}^{T} \beta_i^t X_i^t + \tau_{treatment}(I_i^t - I_i^{t-1}) + \tau_{ant}(J_i^t - J_i^{t-1}) + \epsilon_i^{T,t} \tag{2.2}$$

where I_i^t represents the indicator variable that is equal to 1 if a transaction took place in the treatment period (>2004), and is equal to zero otherwise. J_i^t represents the indicator variable that is equal to 1 if a transaction took place in the anticipation period (2000 — 2004), and is equal to zero otherwise. If more than one sale took place in a treatment time and in a treatment area, $(I_i^t - I_i^{t-1})$ would take the value of zero. In this manner, we consider the treatment effect only once for each property i. $\tau_{treatment}$ is then interpreted as the overall effect of the new transportation development on the price of housing, including observed and unobserved factors. τ_{ant} reflects the effects of "anticipation" for the new highway, as capitalized in the housing transaction prices. The period 1995–1999 is defined as the control period. The second specification we use is focused on identifying the effects of the three main externalities of transportation development observed in this research: improved accessibility, and increased traffic intensity and noise pollution. The model specification below is a modification of equation 2.2 in the sense that three variables are added to describe changes between time periods in the levels of the externalities mentioned above:

$$\begin{aligned}\log \widetilde{P_i^{T,t}} = &\sum_{t=0}^{T} \beta_i^t X_i^t + \mu_1^{T,t} \widetilde{W_{noise}^t} + \mu_2^{T,t} \log \widetilde{W_{acc}^{T,t}} + \mu_3^{T,t} \widetilde{W_{dens}^{T,t}} + \\ &\sum_{t=1}^{3} \sum_{j=1}^{N} \pi_j^t N_j^t + \epsilon_i^{T,t}\end{aligned} \tag{2.3}$$

where $\widetilde{W_{noise}^t} = (W_{noise}^T - W_{noise}^t)$ is a dummy variable that is equal to 1 if a property is located within 300m of the highway. As in the previous model specification, this expression becomes equal to zero if both transactions in period T and t have occurred after the completion of the highways in the summer of 2004. $\widetilde{W_{acc}^{T,t}} = \frac{W_{acc}^T}{W_{acc}^t}$ indicates changes in accessibility levels. $\widetilde{W_{dens}^{T,t}} = (W_{dens}^T - W_{dens}^t)$ is a dummy variable that is equal to 1 if a property is located up to 1 km from a highway interchange. $\mu_1^{T,t}, \mu_2^{T,t}, \mu_3^{T,t}$ are estimated with respect to W_i, and measure the effect of changes in each of the externality levels on the price of housing. As the externality variables take a value different from zero only for treatment transactions, when changes in externality exposure levels occur, the corresponding estimators $\mu_1^{T,t}, \mu_2^{T,t}, \mu_3^{T,t}$ essentially function as DID estimators.

The effects of the externalities of new highway developments may differ over municipalities, depending on unobserved neighborhood characteristics (Martínez and Viegas, 2009). This issue is addressed by including a neighborhood interaction dummy, N_j^t, which indicates that property i is located in municipality $j = 1 \ldots N$, and was sold in period t, before, during or after the completion of the road. The estimator π_j^t is not interpreted, as the latter term of the model is used to account for unobserved time trends in each of the municipalities in the treatment area. Notably, while this approach aims to improve our control for endogeneity followed by the presence of both unobserved time-invariant and time-variant factors, the use of time and municipality dummies may arguably still not fully account for unobserved time trends within micro-locations. These suspicions can be mitigated when considering that the trajectory of the highways was determined in most parts based on existing infrastructure that was built in preceding decades. However, if such trends are indeed present, they have likely affected the local assignment of new highways, which would imply overestimation of our results.

A key problem that arises from equation 2.3 is that the model may not sufficiently identify the differences in the marginal effects of the externalities between neighborhoods. For example, some neighborhoods may be more sensitive to one externality than another. In that case, it may be useful to use more than one estimator to explain the effects of changes in externality level. Therefore, equation 2.3 may be modified using a more flexible specification that is able to accommodate neighborhood differences in the estimated values. This may be done by interacting the neighborhood effects with the externality variables as specified in equation 2.4:

$$\begin{aligned} \log \widetilde{P_i^{T,t}} = & \sum_{t=0}^{T} \beta_i^t X_i^t + \sum_{j=1}^{N} \mu_{1,j}^{T,t} * A_i^{T,t} * \widetilde{W_{noise}^t} + \\ & \sum_{j=1}^{N} \mu_{2,j}^{T,t} * B_i^{T,t} * \log(\widetilde{W_{acc}^{T,t}}) + \sum_{j=1}^{N} \mu_{3,j}^{T,t} * C_i^{T,t} * (\widetilde{W_{dens}^{T,t}}) + \epsilon_i^{T,t} \end{aligned} \tag{2.4}$$

Specification 2.4 aims to estimate individually the effects of the externalities in each of the municipalities in the treatment area. $A_i^{T,t}, B_i^{T,t}, C_i^{T,t}$ are dummy variables that indicate a specific municipality affiliation. Namely, we allow $\mu_{1,j}^{T,t}, \mu_{2,j}^{T,t}, \mu_{3,j}^{T,t}$ to be estimated separately for each municipality, thus permitting spatial flexibility between the estimated effects of each externality.

2.4 Data description and summary statistics

2.4.1 Case studies

To examine the effects of transportation infrastructure development, we use two case studies. The purpose of this is to increase the robustness of the results by increasing the available sample, and also to compare the effects of different highway development projects in relatively similar areas in terms of accessibility and population density. The projects chosen were those of the A30 and the A50 Rijkswegen (highways) in the east of the Netherlands, both completed in the summer of 2004 (see appendix 2.A).

2.4.2 Housing price data

House transaction data were made available through Kadaster, the Dutch land registry. The data used in the research include all housing transactions in five Dutch provinces – Noord-Brabant, Gelderland, Flevoland, Limburg, and Utrecht (excluding the city of Utrecht itself) – for the period 1995–2011. As the data are used exclusively for repeat sales analysis, which requires properties to be sold more than once, we include only housing transactions for properties that were sold twice or more – approximately 438,000 transactions, or 37% of the original data. For the purpose of the control and treatment analysis, three time periods are defined, as follows:

i "Control period" (1995-–1999) : Before highway development began.

ii "Anticipation period"(2000—2004): During highway development works.

iii "Treatment period" (2005—2011) : After highway development ended.

Summary statistics of the housing price data are specified in Table 2.1 and in appendix 2.B.

2.4.3 Accessibility indicators and data

To compare accessibility levels between postal code areas over time, we use the economic potential indicator (see, for instance, Gutiérrez (2001); López et al. (2008); Spiekermann and Neubauer (2002); Vickerman et al. (1999)). The economic potential indicator is originally derived from the gravity model and is formulated as follows:

$$ACC_i = \sum_{j=1}^{N} \frac{P_j}{I_{i,j}^{\alpha}} \tag{2.5}$$

Table 2.1: House transaction prices – summary statistics (Euro, 1995-–2011)

Year	N	Mean	sd	min	max
1995	22,007	97,508	43,784	25,071	726,048
1996	25,928	108,272	50,073	25,185	1,361,341
1997	28,413	120,083	54,611	25,033	1,050,501
1998	30,167	132,365	64,160	25,461	1,815,121
1999	31,818	153,625	80,845	25,412	1,928,566
2000	30,579	177,628	95,485	25,226	1,633,609
2001	31,499	191,645	103,294	25,626	1,953,905
2002	30,414	202,205	104,736	25,000	1,816,028
2003	28,400	208,224	103,064	26,500	1,925,000
2004	27,013	215,623	108,342	25,000	1,930,000
2005	28,434	224,762	113,525	25,000	1,850,000
2006	28,577	235,472	120,459	27,500	1,875,000
2007	26,614	247,515	135,789	27,000	1,937,500
2008	23,723	253,800	139,564	25,000	2,000,000
2009	15,999	240,780	131,068	25,000	2,000,000
2010	15,022	241,178	132,545	29,000	1,950,000
2011	14,207	239,340	132,486	27,000	1,850,000
All	438,814	189,616	113,827	25,000	2,000,000

where ACC_i is the economic potential accessibility of postal code area i, $I_{i,j}$ represents the impedance level, scaled with the power α, and P_j represents the activity factor. In this research, the impedance function was constructed by using the expected road network travel time in seconds between each of the four-digit postal code (PC4) areas in the examined area in the Netherlands. Travel time was calculated from the center of the constructed area for each PC4 area using an open-street geographic information system. The activity element was computed based on population data in each PC4 area, which was available from Statistics Netherlands.[2]

The economic activity potential indicator is also used in this research as it is suitable for the relatively large study area, which varies in terms of level of population density. As commonly used in the literature, we set the distance decay parameter to $\alpha = 1$ (Bruinsma and Rietveld, 1998; Gutiérrez, 2001; Gutiérrez et al., 2010). Table 2.2 describes the changes in accessibility levels attributed to the completion of the new highways in each of the treatment municipalities.

2.4.4 Noise level data

The effects of traffic noise are limited to the local surroundings of the highway and they depend on factors such as traffic density and flows at each hour of the day, natural or planned noise barriers, and others. Therefore, in the absence of exact data measurements, it is difficult to determine the distance from the road up to which noise pollution is apparent. Wilhelmsson (2000) uses a noise model to argue that for distances less than 300m from a road (in suburban areas in Sweden), the marginal contribution of traffic noise to the surrounding noise pollution is substantial. Theebe (2004) finds that noise pollution affects housing prices in the Amsterdam area only at levels above 55 dB. Based on data from the Dutch Ministry of Infrastructure and Environment (RWS), these noise levels exist up to roughly 200m from the A30 and A50 roads, depending on the exact location. As it is difficult to determine the exact limit of the noise effects, we identify areas exposed to noise pollution levels using a distance dummy of 300m from the road. It is worth noting that the A30 and the A50 are located in low population density areas and thus relatively few properties are observed within close distances of the highways (Table 2.3).

[2] In 2004, the year in which the highways were completed, there were 3,993 four-digit postal code areas in the Netherlands, which had between 50-22,860 inhabitants, with a maximum of approximately 11,685 households in the city of Amsterdam. The average PC4 area had 4,070 inhabitants, or 1,765 households per PC4 area.

Table 2.2: Summary of changes in accessibility in the municipalities directly adjacent to the treatment areas

	Municipality	Postal code	Mean level of change in accessibility (%)	Freq.
A30	Barneveld	3771	2.84	913
	Barneveld	3772	3.00	889
	Barneveld	3773	2.34	416
	Voorthuizen	3781	2.19	746
	Ede	6711	2.15	885
	Ede	6712	2.49	549
	Ede	6713	1.70	1,389
	Ede	6714	2.60	832
	Ede	6715	2.04	695
	Ede	6716	2.28	1,340
	Ede	6717	2.55	1,816
	Bennekom	6721	2.70	1,315
	Lunteren	6741	3.12	720
	Total A30		2.44	12,505
A50	Uden	5401	10.40	950
	Uden	5402	10.32	568
	Uden	5403	8.10	1,143
	Uden	5404	10.97	298
	Uden	5406	10.08	224
	Volkel	5408	6.61	98
	Veghel	5463	6.01	733
	Veghel	5464	8.09	117
	Veghel	5465	8.89	20
	Veghel	5466	9.77	35
	Veghel	5467	5.88	1,129
	St Oedenrode	5491	6.72	688
	St Oedenrode	5492	6.27	394
	Son	5691	7.33	925
	Total A50		7.86	7,322

Note: Population growth rates were separated from the calculation of the changes in accessibility values noted here.

Table 2.3: Number of properties which are located within 300m from a new highway, or 1km from a new interchange.

	300m from a new highway (Noise pollution)	Share of total properties in a municipality (%)	1km from a new interchange (Traffic intensity)	Share of total properties in a municipality (%)
Barneveld	10	0.10	67	0.60
Bennekom	-	-	-	-
Ede	1	0.00	108	0.40
Lunteren	-	-	8	0.20
Voorthuizen	-	-	-	-
A30	11		183	
Sint Oedenrode	19	0.30	69	1.00
Son	14	0.30	17	0.40
Uden	54	0.40	207	1.40
Veghel	1	0.00	27	0.30
Volkel	-	-	-	-
A50	88		320	
Total	99		503	

Note: Transactions noted here are only those made in the 3rd period, during or after 2005.

2.4.5 Traffic intensity data

Estimating the negative effects of increased traffic intensity on housing prices poses several challenges. First, increased levels of traffic intensity are closely related to increases in both noise and accessibility levels and can indirectly affect housing prices through these externalities. However, after controlling for changes in noise pollution and accessibility levels, traffic intensity may affect housing prices through its indirect effect on public health, road safety, and other factors (such as air pollution from traffic emissions, or general residential ambience). These effects are particularly local and predominate almost exclusively in densely populated areas. Therefore, we define properties that are located within a distance of 1 km from a new highway interchange as being exposed to changes in traffic intensity levels as we assume that such properties in the immediate proximity of an interchange will be exposed to higher levels of traffic. The limit of 1 km was chosen based on the characteristics of the study area. In the study area, a 1 km distance from an interchange usually includes a limited number of regional roads that connect the highway interchanges to nearby towns. Also, this range is usually sufficiently broad to include a sufficient number of properties, but so large that it includes whole neighborhoods or overlaps areas influenced by other interchanges (see appendix 2.C).

Table 2.3 contains summary statistic for properties that are located in proximity of 1km to a new A30 or A50 interchange.

2.4.6 Treatment and control areas

The treatment area is defined based on changes in accessibility level. Postal code areas that have experienced an increase of over 2.5% in accessibility level are defined as treatment areas. In our chosen study area, the five provinces in the south and west of the Netherlands, approximately 3% of postcodes meet this criterion. In relation to control area definitions, we use two scenarios based on aerial distance from the road or changes in accessibility levels. In the first scenario, referred to as the “distance study area,” we arbitrarily define the control area as postal code areas that are located within 10 km of the development projects, excluding the city of Eindhoven (see appendix 2.D). This definition is motivated by the assumption that postal code areas within this range share common spatial and housing market characteristics with the treatment area, while not all are influenced by the development of the highways. In the second scenario, referred to as the ”accessibility study area”, we define the control area as postal code areas that have experienced over 1.5% but below 2.5% change in accessibility level (see appendix 2.D). In the study area, approximately 20% of postal code areas meet this criterion. The advantage of using such an extensive control area is that it has a larger observation base of changes in accessibility levels. However, this definition may include areas that are too distant and different from the treatment area, which may result in a biased estimated effect. One notable disadvantage of both scenarios is that most postal code areas in the municipality of Ede are not included in the treatment area as they have experienced an accessibility level increase of less than 2.5% (see Table 2.2), but they are included in both control area scenarios. As Ede is located at the southern entrance of the A30 and is situated between two ”treatment” municipalities, it can be argued that it cannot serve as an appropriate control area. However, excluding Ede from either control area did not cause a significant change in the DID coefficient values or in statistical significance levels, and thus it was left in the study as a control area.

2.5 Estimation results

2.5.1 Repeat sales and difference-in-difference (DID) estimators

The first specification to be estimated is the repeat sales model including the DID estimators, as specified in equation 2.2. The purpose of this specification is to address the question of how the development of the A30 and the A50 highways has affected housing prices in the nearby municipalities (see results in Table 2.4). The model was estimated using different DID estimators. The specifications reported in columns 1 and 2 include DID estimators for the treatment and anticipation effects for both highways under each treatment control area scenario. The specifications in column 3 and 4 include DID estimators for the treatment and anticipation effects for each highway individually under both treatment control area scenarios.

Table 2.4: Estimation results for the repeat sales and DID estimators

	Distance study area (1)	Accessibility study area (2)	Distance study area (3)	Accessibility study area (4)
Anticipation (total)	0.0471***	0.0532***		
	(0.0124)	(0.0122)		
Treatment (total)	0.0251***	0.0431***		
	(0.00497)	(0.00446)		
Anticipation (A30)			−0.0419***	−0.0358***
			(0.00818)	(0.00677)
Treatment (A30)			−0.0265***	−0.00876*
			(0.00515)	(0.00470)
Anticipation (A50)			0.0869***	0.0929***
			(0.0170)	(0.0170)
Treatment (A50)			0.0489***	0.0670***
			(0.00618)	(0.00574)
Year dummies	Yes	Yes	Yes	Yes
Observations	42,148	91,613	42,148	91,613
R^2	0.826	0.835	0.827	0.836

Notes: (i) Explained variable: log of price differences, (ii) Standard errors in parentheses, (iii) *$p<0.1$, **$p<0.05$, ***$p<0.01$

The results reveal several interesting findings. A notable finding is that almost all DID estimators are positive, significant, and their values remain relatively constant between each specification, and each treatment control area scenario. The fact that the coefficients and statistical significance levels vary relatively little between different definitions of control area indicates that the estimated results are quite robust. Combining the treatment in both highways (columns 1 and 2), we see that the effect of the new highways has increased housing values in the surrounding residential area by approximately 2.5–4.3%. Moreover, prices increased by an even higher rate of approximately 5% before the highways were completed. This finding matches the assumption that the residents in the treatment area attributed a positive value to the development of the highways and this is capitalized in the value of the residential properties. Second, the value of the DID estimator for the anticipation period is much higher than the value of the DID estimator for the treatment period. Two possible explanations can be provided for these results. First, this may suggest that the housing market in the treatment area anticipated the change in prices and began adjusting long before the roads were completed in 2004. Second, it may also be that before the completion of the highways, the public expected that the effects would mostly be positive (probably due to increased accessibility). However, after the completion of the projects, the negative effects of increased traffic and noise pollution would have become apparent, and this is reflected by the lower values of the post-treatment DID estimators.

The results for the A50 (columns 3 and 4) match the expectation that houses in the treatment area would increase in value. Compared to both the distance and accessibility-based control areas, houses that were sold during the four years prior to the completion of the A50 gained approximately 9% in value. After completion, the A50 added 4.9–6.7% to the values of the houses in the nearby municipalities.

However, the estimated results for the A30 show an opposite and unexpected pattern. Compared to properties in the distance-based control area (within 10 km of the A30), houses in the municipalities along the A30 have not experienced an increase in value and the effect is estimated to be negative. Moreover, in the four years prior to the completion of the road, housing transactions were approximately 4% lower in value, implying a negative anticipation effect. A possible explanation for these results is that the populated municipalities at the edges of the A30 were already relatively accessible before the construction of the A30 (Barneveld is located close to the A1 between Amersfoort and Apeldoorn, and Ede is located close to the A12 between Utrecht and Arnhem). In fact, compared to the relatively high rates of improvement in accessibility (approximately 5.9–10.9%) experienced in the A50 municipalities, the development of the highway improved accessibility levels only slightly more than 2.5% in the A30 area (see Table 2.2). This value is not notably different from the increase in accessibility in both control areas. This suggests that residents of these municipalities have not enjoyed increased accessibility from the construction of the A30 and perhaps the dominant effects in this area are those of increased traffic and noise pollution following the development of the highway.

2.5.2 Repeat sales: including externality levels

The second model specification includes direct estimation of the effects of each of the three externalities examined on the price of housing. The specification is as described in equation 2.3 and addresses the question of how home owners and residents of the region's municipalities value each of the highway externalities imposed by the development of the A30 and the A50 (see Table 2.5). Due to network effects, the development of the highways resulted in changes in accessibility levels throughout the study area. Therefore, this specification no longer makes use of the restricted treatment and control areas used previously. We begin by including only the positive accessibility effects in the model (Table 5, column 1). The reason for this is that a significantly larger number of properties are assumed to be influenced by accessibility effects compared to negative noise and traffic intensity effects. The negative effects are added later and are specified in column 2.

Table 2.5: Estimation results from the repeat sales model: including externality levels

	(1)	(2)
$\leq$ 300 m from the highway (noise)		−0.0360*
		(0.0218)
Log. changes in accessibility levels	1.760***	1.765***
	(0.0780)	(0.0782)
$\leq$ 1 km from interchange (traffic density)		−0.0298**
		(0.0129)
Year dummies	Yes	Yes
Municipality-time interaction dummies	Yes	Yes
Observations	269,129	269,129
R^2	0.820	0.820

Notes: (i) Explained variable: log of price differences, (ii) Standard errors in parentheses, (iii) *$p<0.1$, **$p<0.05$, ***$p<0.01$

As expected, the coefficients of the changes in accessibility levels show strong positive values. The estimated elasticity between transaction price ratios and changes in accessibility levels is approximately 1.76, which means that a change of 1% in accessibility levels is expected to result in a 1.76% increase in the transaction price ratio. The estimated value is robust and changes little when negative externalities are introduced in the model (column 2). As the increase in accessibility level in the municipalities along the A50 ranges from

5.9% to 10.9% (see Table 2.2), this increase translates to an approximate increase of 10–18% in the transaction price ratio. The estimated coefficient for changes in traffic density levels is also according to expectations and its value is estimated at approximately -0.0298, implying that properties located within 1 km of a new interchange have also experienced a decline of approximately 3% in value. The noise pollution coefficient is also estimated to be negative and statistically significant at -0.0360, indicating that properties in this range have experienced a decline of 3.6% in value.

The strong positive value of the accessibility coefficient indicates that for many properties, the improvement in accessibility is the dominant effect resulting from the development of the highways. Moreover, as some properties are exposed to multiple externalities, the positive accessibility effects can often offset the negative effects. This corresponds to the results of the single DID estimator specification (Table 2.4), in which the combined effect of the development project is found to be positive and significant.

2.5.3 Quasi-rents and marginal willingness-to-pay (WTP) analysis

The results of this specification can also be used to place monetary value on the development of the new highways. Following Mohring (1965), investment in transportation infrastructure development can be analyzed by the quasi-rents generated by the development project. This logic was applied by Klaiber and Smith (2010), who measured changes in capitalized housing prices to estimate the quasi-rents following highway segment additions. Focusing on the accessibility coefficient (Table 2.5, column 1), we are able to conduct a similar analysis and calculate households' WTP for accessibility improvement, which is defined as the change in housing price that would keep utility constant with any change in accessibility.

As changes in accessibility levels differ between postcodes, we begin by calculating the quasi-rents for the average property in each PC4 area. To obtain this, we multiply the estimated elasticity between price and accessibility with the average housing price in each PC4 area (in 2004, when the highways were completed), and the percentage of accessibility improvement that the area had experienced. Aggregated values are then calculated by multiplying the WTP values by the total housing supply in each PC4 area. The aggregation is undertaken under the assumption that all houses in a certain postal code area have experienced the same price effect due to the change in accessibility levels. As accessibility levels have increased relatively little in the municipalities along the A30, this analysis focuses on the A50 area.

In the municipalities located along the A50, significant changes in accessi-

Table 2.6: Average housing prices per PC4 in the A50 surrounding postal codes areas, and WTP for accessibility values

Municipality	Postcode	Accessibility change (%)	Mean price (€thousand)	Average WTP (€thousand)	Housing supply	Aggregate WTP (€thousand)
Uden	5401	10.4%	200.0	36.60	4,235	154,987
Uden	5402	10.3%	218.7	39.74	3,220	127,948
Uden	5403	8.1%	221.3	31.54	3,785	119,378
Uden	5404	11.0%	338.5	65.36	1,685	110,134
Uden	5406	10.1%	275.2	48.80	1,920	93,700
Volkel	5408	6.6%	257.8	29.99	1,305	39,140
Veghel	5463	6.0%	185.7	19.63	2,805	55,074
Veghel	5464	8.1%	281.7	40.12	2,130	85,460
Veghel	5465	8.9%	320.0	50.04	605	30,276
Veghel	5466	9.8%	384.6	66.12	565	37,356
Veghel	5467	5.88%	204.8	21.20	3,100	65,730
St Oedenrode	5491	6.7%	230.1	27.23	4,000	108,908
St Oedenrode	5492	6.3%	235.0	25.93	3,085	79,997
Son	5691	7.3%	307.0	39.63	4,280	169,618
Total					36,720	1,277,707
WTP per km of highway (€thousands)						37,580

bility levels are observed. This corresponds with high WTP values per postal code in this area (see Table 6), which are found to range between €19,000 and €66,000. These values reflect a price increase equivalent to approximately 10–18% of the housing price. Examination of the aggregated WTP values reveals that the region's benefits from the highway development are estimated to be approximately €1,277 million. When the length of the A50 is taken into consideration, the benefits from the projects become comparable to other previous findings. The length of the developed A50 segment is 34 km and therefore the benefits per developed km are equivalent to approximately €37.5 million. These values show some similarity with those found by Klaiber and Smith (2010), which were estimated at US $73 million to US $273 million per mile (or approximately €30 million to €125 million per km of additional highway).

2.5.4 Repeat sales: including externality levels and neighborhood effects

As municipalities and neighborhoods are assumed to be heterogeneous, it is possible that the marginal effects of each of the externalities are also different between them. The purpose of the model that is specified in equation 2.4 is to check for differences in neighborhood effects by creating interaction variables between each of the 30 municipalities in the treatment area and the changes in levels of externalities (see Table 2.7).

It should be highlighted that as the number of available observations is relatively low, a biased estimator is probable in many municipalities. Also, some municipalities in the treatment area are not exposed to the negative highway externalities at all. Only four municipalities include multiple properties within 300m of the highway (Barneveld, Son, Sint Oedenrode, and Uden) and seven municipalities include properties within 1 km of a new interchange (Barneveld, Ede, Lunteren, Son, Sint Oedenrode, Uden, and Veghel). This makes it difficult to undertake a proper comparison of the marginal effects of traffic and noise pollution levels on the prices of houses in different municipalities.

As expected, the municipality-specific coefficients for changes in accessibility are positive and significant for the A50 municipalities, and their values vary between 0.53 and 1.41, which fits with the results of the previous estimation. The estimated values show that accessibility elasticities are different in each municipality, probably due to unobserved heterogeneity in preferences or spatial characteristics in each town. As mentioned above, it is possible that some of the results may be biased due to the low number of observations. Interestingly, the values of accessibility coefficients in the A50 area increase from the south of the A50 (Son) to the north (Volkel). This may suggest that although A50 municipalities have experienced different levels of improvement in accessibility from the highway network perspective, the effects are valued higher in towns in which accessibility to Eindhoven has been improved in particular, most likely as it is the largest employment center in the province.

Table 2.7: Estimation results from the RS model: Including externalities and neighborhood effects per municipality

			(1)
Log. changes in accessibility levels	A30	Barneveld	−0.157
			(0.142)
		Bennekom	0.130
			(0.189)
		Lunteren	0.527*
			(0.291)
	A50	Son	0.531***
			(0.0736)
		Sint Oedenrode	0.553***
			(0.125)
		Veghel	0.682***
			(0.160)
		Uden	1.139***
			(0.218)
		Volkel	1.406***
			(0.372)
$\leq$ 300m from the highway (noise)	A50	Uden	−0.0602**
			(0.0301)
$\leq$ 1km from interchange (traffic density)	A50	Uden	−0.0542*
			(0.0329)
Year dummies			Yes
Control area municipalities interaction variables			Yes
Observations			269,129
R^2			0.818

Notes: (i) Explained variable: log of price differences, (ii) Standard errors in parentheses, (iii) *p<0.1, **p<0.05, ***p<0.01

Among the A30 municipalities, the estimated values in the municipalities of Barneveld and Bennekom are statistically insignificant, which supports the argument presented earlier that the development of the A30 has not significantly improved accessibility in these municipalities, and that it is likely rather to be correlated with increased exposure to negative externalities. Unlike Barneveld and Bennekom, the town of Lunteren presents a positive and significant coefficient at the 10% level. This might be explained by the location of Lunteren. As Lunteren is located at the center of the A30 and was not previously connected to a major highway, it is possible that the inhabitants of Lunteren attached positive values to the improvement in accessibility that resulted from the development of the A30, unlike those of other municipalities in the area.

Due to the lack of observations, the only municipality in which an interpretation of the noise pollution coefficient can be provided is Uden, where the noise coefficient is estimated based on 54 observations (see Table 5). The results in Table 2.5 show that the estimated value of the noise coefficient is significant and negative at -6.0%, indicating that properties at this proximity to the A50 in Uden are valued negatively compared to the rest of the sample. Uden is also the only town in which a sufficient number of transactions occurred within 1 km of a new A50 interchange and thus the traffic density effect can be estimated. As expected, the estimated value is negative at -5.4%, statistically significant at the 10% level.

The results of the model specifications provide several important findings. The first model confirms that the total effect of highway development on the price of housing is not only positive, but also salient, and may even be stronger during the period before the project ends. The implicit conclusions are that for most properties in the vicinity of the highway, the positive externalities are more dominant than the negative ones, and these positive effects on the local housing markets are capitalized long before the end of construction. The second model separates the effect of the development of the highways into the three major externalities examined and estimates the value of each of these from the perspective of the public. The results show a positive valuation for improved accessibility, and aversion related to noise pollution and traffic intensity. Finally, the results of the specification that addresses the issue of neighborhood heterogeneity in relation to externality effects show that price-accessibility elasticities are not constant. Although mostly positive, the values of elasticities may depend on unobserved heterogeneity between the spatial characteristics of the towns or the preferences of the inhabitants in each town.

2.6 Conclusions and Discussion

In this chapter, we have explored the effects of transportation infrastructure development on the price of housing. To estimate these effects, we apply three different specifications of the repeat sales model. The availability of extensive information on residential property transactions allowed the use of repeated sales and difference-in-differences to study the effect of new highways in great detail. In the estimation process, we control for each highway externality separately, neighborhood effects and anticipation effects. The findings support our expectations. Improved accessibility is found to be valued positively, whereas noise and traffic intensity are found to affect housing prices negatively. In most properties in the treatment area, the combined effect of the three externalities is found to be positive, as reflected in the overall average increase in the transaction prices of houses. This implies that the positive value attributed to improved accessibility levels is greater than the negative values attributed to increased noise pollution and traffic intensity levels.

The results obtained in this research may be used primarily in a wider social cost benefit analysis to evaluate whether the development of the roads has been successful from a social perspective and whether the two projects have increased social welfare in their nearby municipalities. As mentioned above, changes in the price of housing reflect the residents' and homeowners' monetary valuation of a newly developed project and therefore they are a key issue in determining changes in social welfare that result from it. However, it is worth noting that although the results presented here are important in determining whether a highway development has been successful from a social perspective, certain other issues should also be considered when addressing this matter. As demonstrated in the case of the A30, negative and insignificant results may appear to imply negative evaluations of the highway. However, the project cannot be regarded as socially unsuccessful only based on results that relate to its immediate surroundings. This highway is likely to have been beneficial in other ways, such as relieving traffic congestion between the east of the Netherlands and the large labor market in the Randstad.[3] A network improvement in accessibility and traffic congestion is unlikely to be reflected in the price of housing in remote areas as not all inhabitants in more distant locations participate in the Randstad labor market and thus the effects on housing prices are expected to be less noticeable. This means that future research should ideally consider wider transportation network effects in the assessment process of changes in social welfare.

The results reported here may also be useful in regional policy. For example, the estimation results and willingness-to-pay (WTP) values may be used to support value setting for future similar highway developments in nearby areas. It is likely that for those in close geographic proximity, spatial and socio-demographic factors are relatively similar and the valuation of the three

[3]Considering the fundamental law of road congestion (Duranton and Turner, 2011), reduction in traffic congestion will arguably only be apparent at the short-term.

different externalities will resemble the estimated values obtained from the towns in our study area. In this respect, future research may examine differences in willingness-to-pay over different population groups, by considering heterogeneity in household characteristics (Bajari and Kahn, 2005; Bajari and Benkard, 2005). Given data availability, such analysis can also shed light on the sorting process that is initiated by the development of the new highways, and to track the demographic change in neighborhoods which may result from this process (Bayer et al., 2004a; Bayer and Timmins, 2005).

Moreover, our results may also be used as a reference for a second-best compensation mechanism, which may be applied to maximize the social gain from the development project. For example, specific taxes can be levied in areas with high increases in property prices to help finance the development project or to provide additional infrastructure investments. Such additional investments may include noise barriers, improvement of traffic conditions near interchanges, or directing compensation to the inhabitants and homeowners in areas where the development project causes a reduction in house prices and social welfare. Such a solution minimizes the welfare loss in areas that do not benefit from the development project and therefore might improve social welfare in these areas.[4]

[4]It is noteworthy that second-best policies should also consider the negative effect on welfare and the deadweight loss associated with increased taxes.

2.A Study area and highway development plan details

The study area includes the five south-eastern provinces of the Netherlands: Limburg, Noord-Brabant, Gelderland, Flevoland, and Utrecht (excluding the city of Utrecht itself). The description of the highway-specific treatment areas is further elaborated under each highway's project description.

A50 (Eindhoven–Oss)

Rijksweg A50 is a highway that stretches 151 km between Emmeloord and Eindhoven, in the east of the Netherlands (see figure A2). Its southern 34 km stretch, between Oss and Eindhoven, in the province of Noord-Brabant, was constructed in the period 2000—2006. The southern part between Eindhoven and Uden was constructed first as a new road that bypasses the small towns north of Eindhoven rather than running through the town centers. The southern part of the project was completed in June 2004. The treatment area for the A50 includes house transactions that took place after the year 2004 in the municipalities of Uden, Veghel, Sint Oedenrode, Son, and Volkel, all of which are adjacent to the A50, or located at a maximum distance of 4 km from it.

A30 (Barneveld—Maanderbroek)

Rijksweg A30 is a highway that stretches 18 km between Barneveld and Maanderbroek (Ede), in the Dutch province of Gelderland (see figure A3). The northern 10 km stretch of the highway was initially opened in 1970 as a 2x2 road between Barneveld and Lunteren. The southern part was also opened in 1970, but consisted of a small 2x1 road that connected Lunteren and Ede Noord. In the period 2003-–2004, the highway underwent a development work that included constructing a final stretch to Maanderbroek and adding a southern access from Rijksweg A12, as well as broadening the highway to 2x2 lanes throughout its length. The construction works were completed in August 2004.

Based on the development project's geographic area and time schedule, the definition for the treatment area for the A30 includes house transactions that took place after the year 2004 in the municipalities of Ede, Barneveld, Lunteren, and Bennekom, all of which are adjacent to the A30, or located at a maximum distance of 5 km from it.

The highways development process is indicated in Table 2.A.1.

Table 2.A.1: Opening dates and lane configuration

Opening dates	From	To	Length	Open
A50	Sint Oedenrode	Veghel	7 km	4-7-2003
(Eindhoven—Uden)	Veghel	Uden–Noord	10 km	8-9-2003
	Ekkersrijt (Eindhoven)	Nijnsel (Sint Oedenrode)	6 km	11-9-2003
	Nijnsel (Sint Oedenrode)	Sint Oedenrode	2 km	4-6-2004
A30	A1 (Barneveld)	Lunteren	10 km	1972
(A1-–Maanderbroek)	Lunteren	Ede-–Kernhem	5 km	4-10-2003
	Ede–Kernhem	A12 (Maanderbroek)	2 km	12-7-2004
Lane configuration	**From**	**To**	**Length**	**Configuration**
A50 (Eindhoven—Uden)	Ekkersrijt (Eindhoven)	Paalgraven (Oss)	34 km	2x2
A30 (A1-–Maanderbroek)	Maanderbroek	Barneveld	18 km	2x2

Note: The northern development of the A50 (Additional 9 km between Uden–Noord and Oss) was opened in stages during 2005 and 2006.

Figure 2.A.1: Map of the study area

Figure 2.A.2: Map of the A50 area

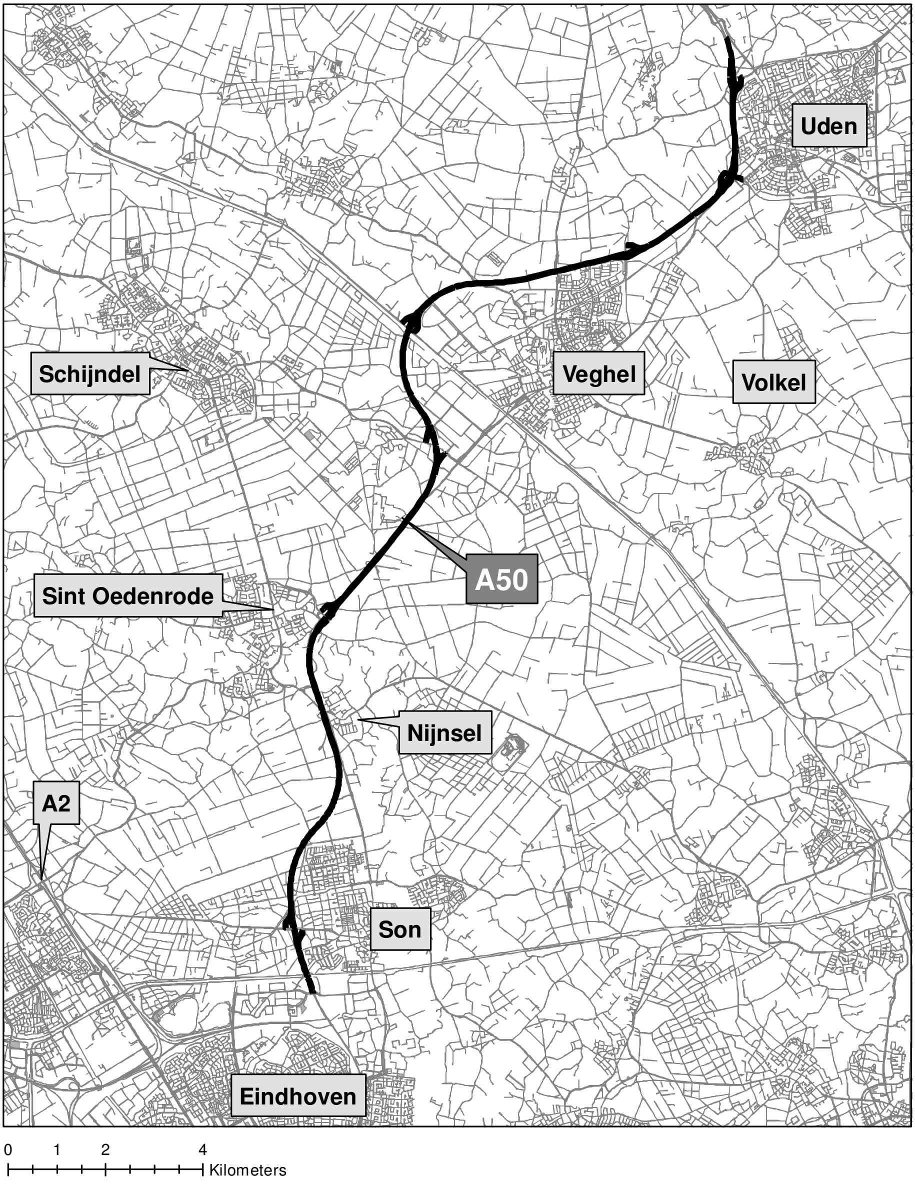
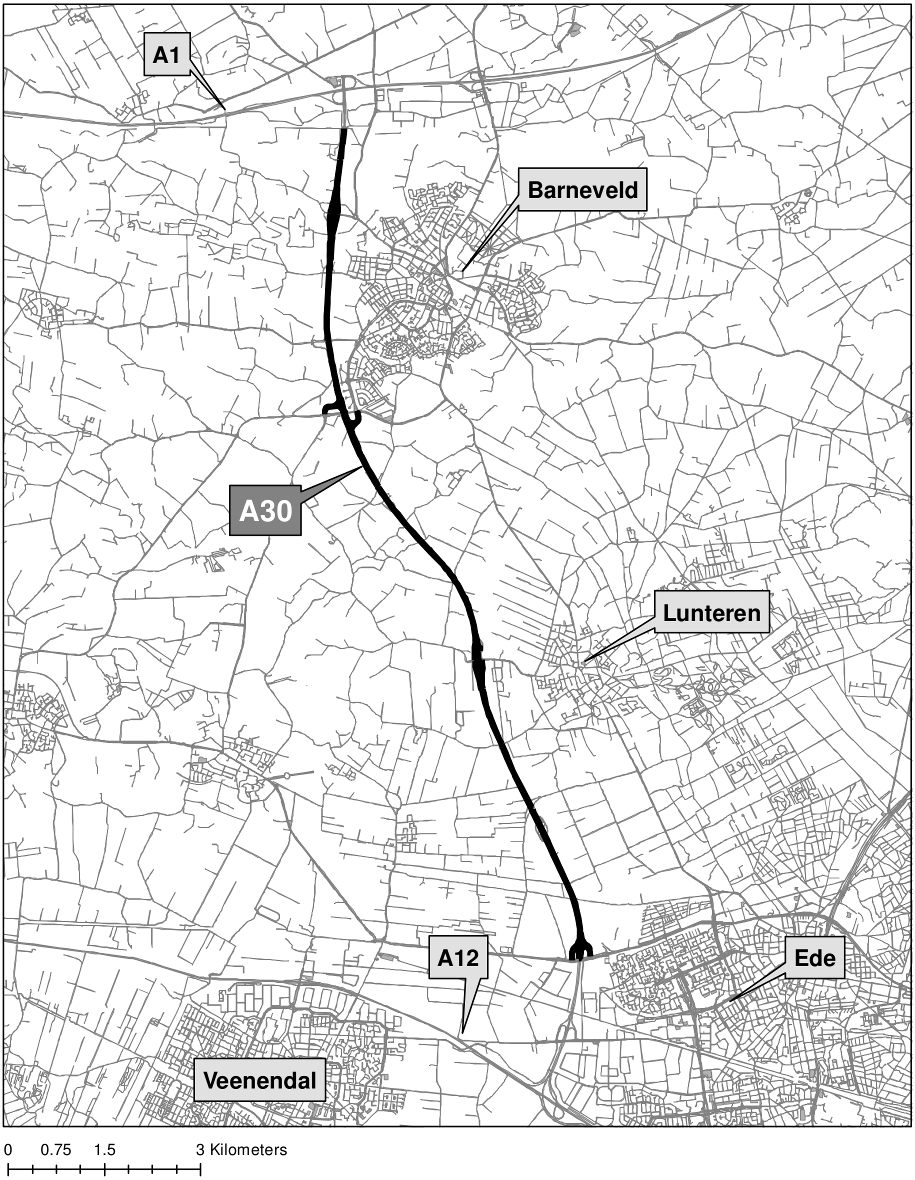


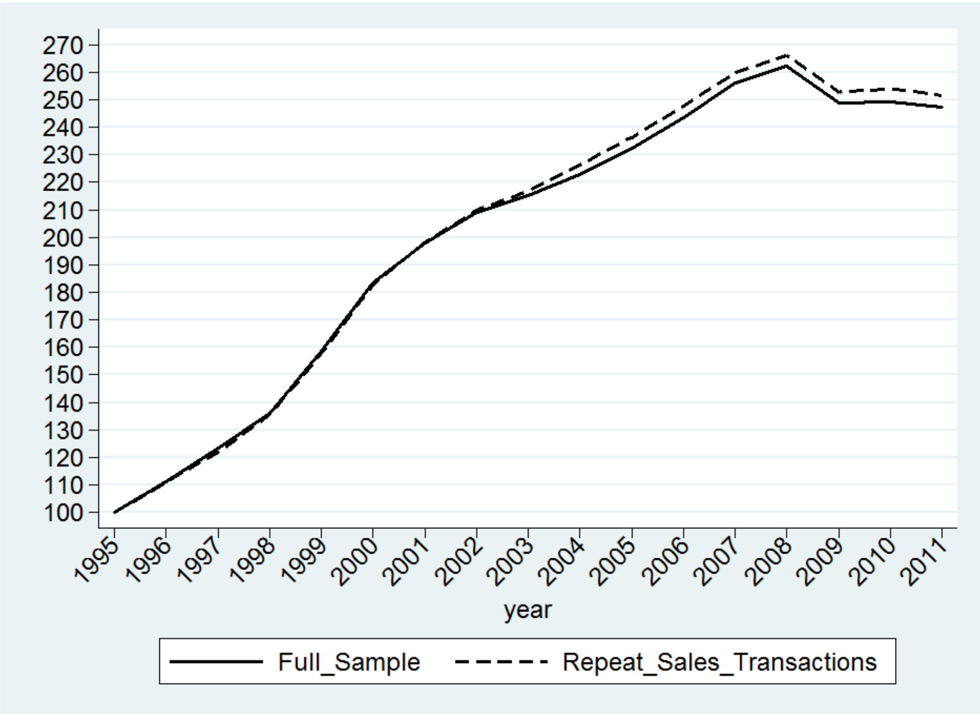
Figure 2.A.3: Map of the A30 area



2.B Price indexes

Price indexes of the study area are based on the full sample (solid line) and on the transactions which were valid for repeat sales analysis (dashed line). The similarity between the two indexes shows that repeat-sales transactions do not demonstrate differences in trends of transactions values, and therefore reduces suspicions of a sample selection bias.

Figure 2.B.1: Price indexes of the study area



2.C Changes in levels of accessibility and regional exposure to externalities

The maps in figures 2.C.1 and 2.C.2 show a graphic depiction of the differences in accessibility levels in each municipality before and after the completion of the A30 and the A50. The accessibility scores were computed using the economic potential activity indicator with the parameter $\alpha = 1$. The map clearly shows how the municipalities along the A50 and A30 have experienced the sharpest increases in accessibility compared to the rest of the Netherlands. Figure 2.C.1 shows the exposure to negative externalities in properties in the municipality of Uden. Properties were considered to be exposed to changes in levels of noise pollution within the vicinity of 300 m from the new A50 and to traffic density within 1 km of a new interchange. Changes in accessibility levels are not shown on this map, but are specified in Figure 2.C.2.

Figure 2.C.1: Areas affected by negative externalities (Uden)

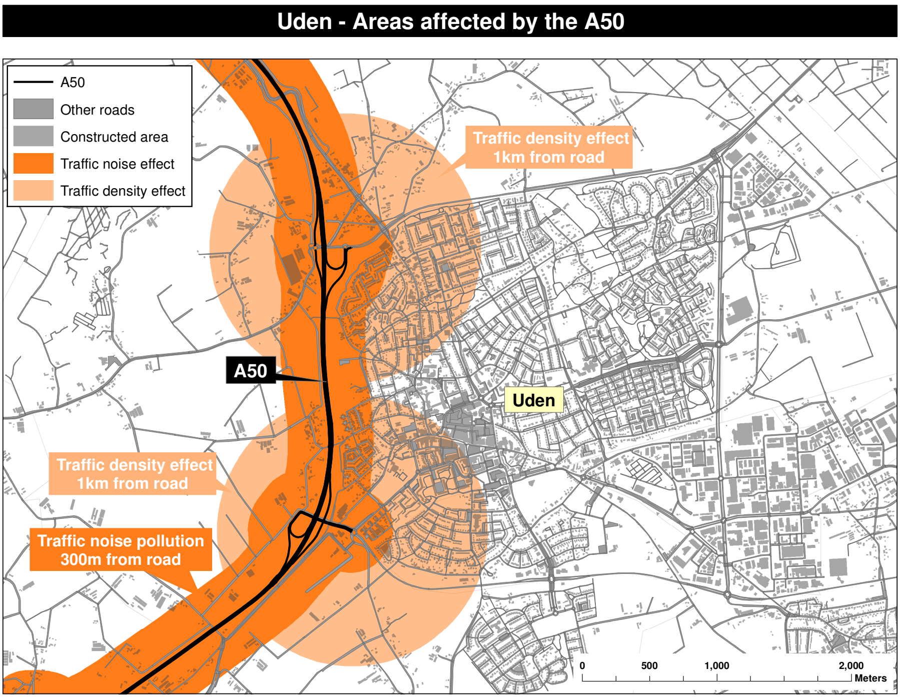
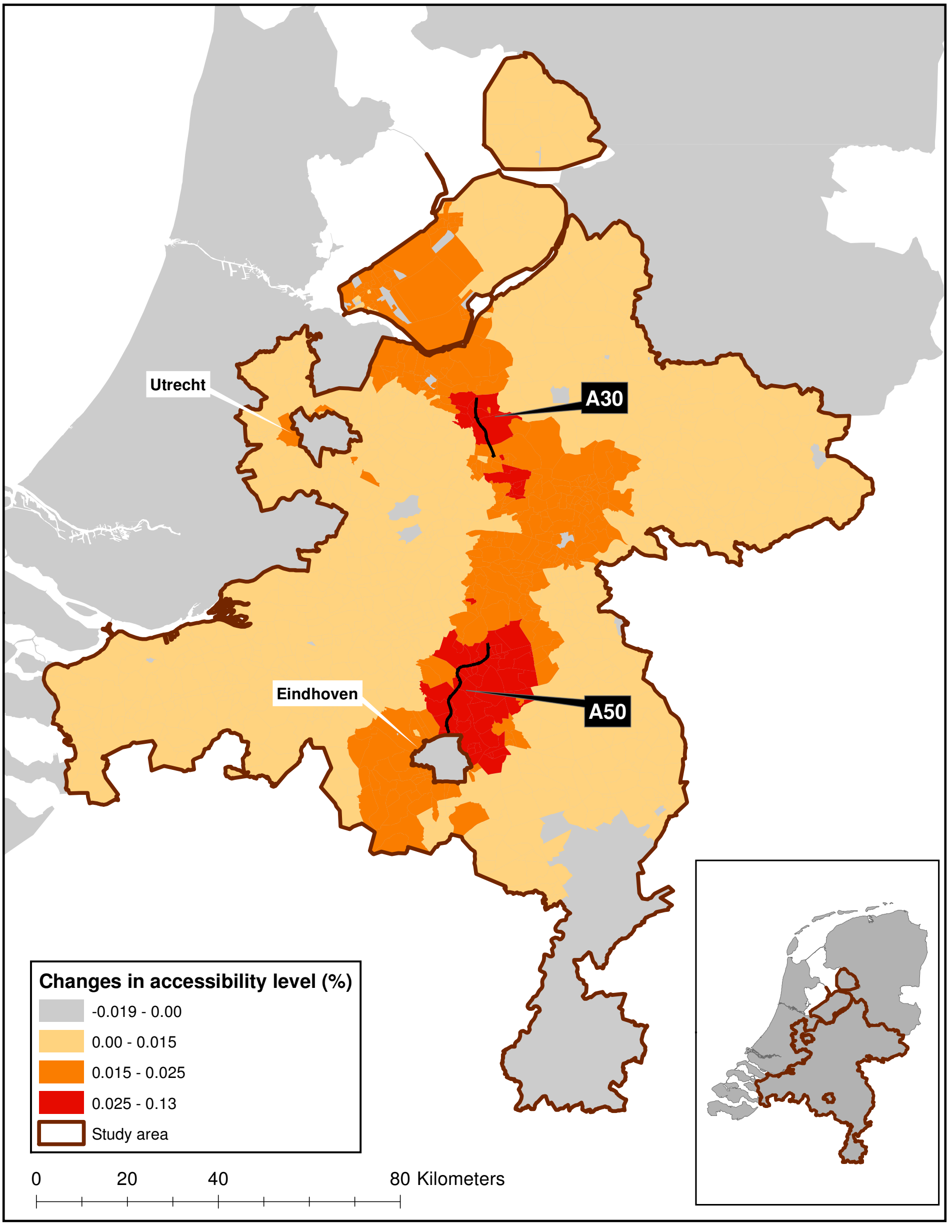
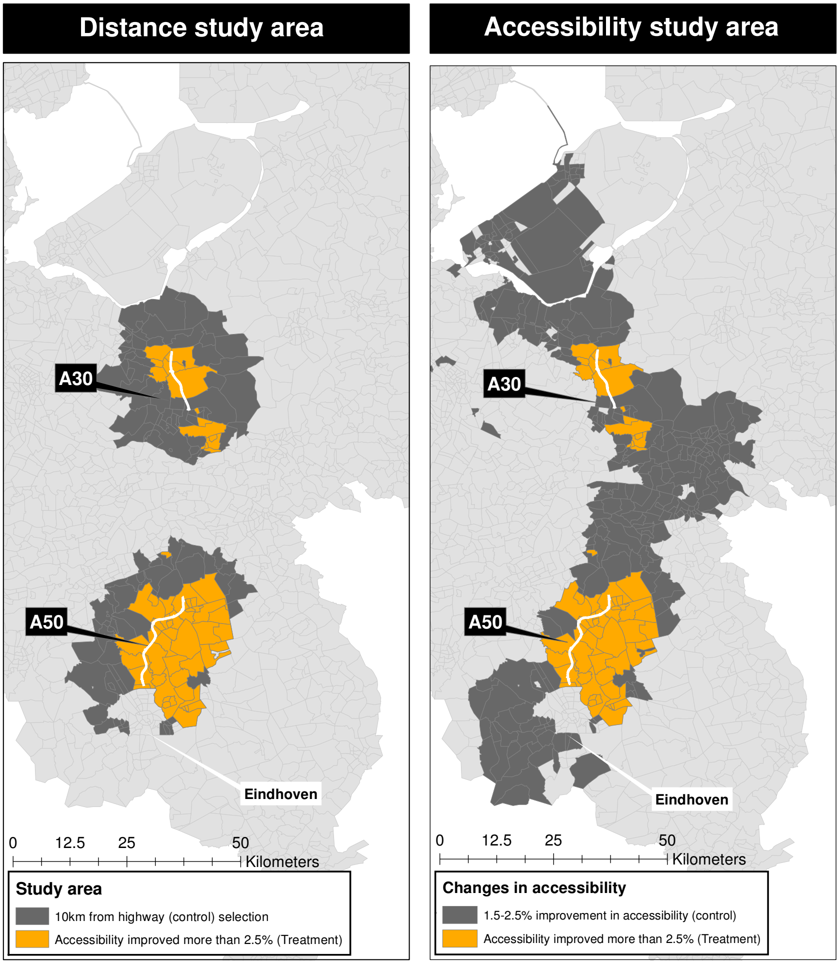


Figure 2.C.2: Changes in levels of accessibility due to the development of the A30 and A50 highways



2.D Difference-in-differences (DID) model study area scenarios

Figure 2.D.1: Maps of the study area scenarios



Chapter 3

The impact of highways on population redistribution: The role of land development restrictions

3.1 Introduction[1]

New highways transform the spatial structure of cities and regions by reducing the costs of commuting to employment centers and improving accessibility in peripheral areas. There is substantial evidence that highways induce suburbanization, reduce population densities in central cities and increase population levels and economic performance of peripheral areas (Baum-Snow, 2007a,b; Chandra and Thompson, 2000; Duranton and Turner, 2012; Garcia-López et al., 2015). However, these predictions may change considerably where urban sprawl is bounded by land development restrictions, prevalent in many cities around the world.

Land development restrictions are often introduced to mitigate urban sprawl. The green belt surrounding London is a well-known example, but similar development restrictions exist in the surroundings of many other cities.[2] Our focus

[1]This chapter is based on joint work with Jan Rouwendal and Jos van Ommeren.
I would like to thank Eric Koomen, Hans Koster and Miquel-Angel Garcia-Lopez for their useful comments and advices. Financial support from ERC Advanced Grant OPTION (#246969) is gratefully acknowledged. I would also like to thank Changjoo Kim and the audience of the paper presentation in the 63rd North American Regional Science Conference for a constructive discussion on the paper.

[2]Notable examples include the Ontario and Ottawa green belts, the São Paolo Biosphere

is on the Netherlands, where the protection of the inner part of the Randstad, a metropolitan area which contains the four largest cities, against urbanization (through the so-called 'Green Heart') is an important aspect of the planning system. In addition, buffer zones contiguous to these cities were introduced to ensure proximity to open space and prevent the merging of urbanized areas.

Not much is known about the effect of urban planning measures on the spatial distribution of the population. There is abundant evidence that land development restrictions increase house prices (Glaeser et al., 2006; Kok et al., 2014), and can impose limitations on employment growth (Hsieh and Moretti, 2017). This implies that restrictions are relevant and that restrictions on land use in the vicinity of large cities induce people to reside further away.

Following the work of Baum-Snow (2007a) it is well known that highway construction was responsible for a substantial share of the suburbanization of US cities. Most of this suburbanization took place through gradual expansion of the urbanized area and increasing densification within the municipal boundaries (Burchfield et al., 2006). In this chapter, we examine how the growth of the highway network in the Netherlands interacted with land development restrictions to transform the spatial distribution of population. More specifically, we address the question whether land development restrictions resulted in different population growth rates than could otherwise be expected on the basis of the expanding highway network.

To determine the causal effect of highway network developments on suburbanization, Baum-Snow used an instrument based on historical plans for network development. Later studies have used similar historical instruments to study, for instance, the effect of highway networks on Chinese cities (Baum-Snow et al., 2017), on land conversion in Spain and on urban structure in Barcelona (Garcia-López et al., 2014, 2015), on innovation in US regions (Agrawal et al., 2017) and on employment levels in Italian cities (Percoco, 2016).

We focus on a large scale expansion of the Dutch highway network in the 1960's, and study its impact on municipal population growth during the decade that followed. We account for endogeneity issues by employing an instrumental variable approach using the 1821 road network. We find that highways have contributed to population growth in peripheral municipalities, but not in suburbs or central cities. Development restrictions resulted in a 'leapfrog' of suburbanization over areas in which development is restricted, population growth primarily occurred in peripheral towns. We also find that the effect of highways on population growth is temporal, as highways seem to have limited effect on population growth more than ten years after the major

reserve and the Seoul green belt (Jun and Hur, 2001).

expansion is completed.

In our empirical strategy, we pay special attention to the treatment of endogenous interaction variables. We use an innovative econometric approach in which we use a single instrument and estimate a single first-stage regression to address multiple interaction terms between an endogenous and exogenous variables. Initially suggested by Balli and Sorensen (2013), we further develop the single first-stage approach and formally show that when the exogenous interaction variable is a categorical variable (which splits the sample into exclusive subsamples), it produces more efficient results than the commonly-applied multiple first-stage approach, under certain testable conditions. Furthermore, we will point out that this approach is biased, in our context at least, due to weak instruments.

The chapter is organized as follows. Section 3.2 provides a theoretical background. Section 3.3 discusses important aspects of Dutch land use planning and road network development. Section 3.4 describes the data. Section 3.5 describes our methodological approach. Section 3.6 includes the estimation results. Section 3.7 includes additional sensitivity analysis and examines long-term effects. Section 3.8 concludes.

3.2 Theoretical background

Following the monocentric model, the development of road infrastructure has a positive effect on city size and suburbanization (Alonso, 1964; Muth, 1969; Mills, 1967). In the closed city version of the model, the rent gradient flattens with a decrease in transportation costs, land in the center becomes less expensive and the city expands through larger average lot sizes. The open city version of the model also predicts a flatter rent gradient, and a larger population. The Roback (1982) model suggests a positive effect of lower local transportation costs on employment on top of that. Glaeser and Kahn (2004) argued that reduction in commuting costs, which accompanied the widespread use of car transport, is the most important driver of suburbanization. This argument was also affirmed by others, notably by Burchfield et al. (2006).

Anas and Moses (1979) and Baum-Snow (2007b) considered an extension of the conventional monocentric model in which space is homogeneous by considering the impact of radial highways on which transport costs are lower. The result is a star-shaped city where each of the 'fingers' grows along a highway. Baum-Snow (2007a) was the first to provide convincing empirical evidence that highways caused suburbanization, as predicted by this model. Later studies have considered the impact of highways on the growth of metropolitan areas and employment (Duranton and Turner, 2012), as well as on increasing economic performance in peripheral areas (Banerjee et al., 2012; Chandra and Thompson, 2000; Michaels, 2008).

The analysis of the effects of highways on population size in peripheral areas has been studied using a different identification strategy, according to which the connection to a highway network of a peripheral town is regarded as an unintended consequence of its location between larger urban areas. Therefore, highway assignment was considered as random within the set of possible trajectories that connect the largest cities (Banerjee et al., 2012; Chandra and Thompson, 2000; Fajgelbaum and Redding, 2014; Michaels, 2008).

We follow the literature and assume a monocentric model setting of a closed city, in which homogeneous workers commute to the city center where employment is concentrated.[3] Building height and residential location are chosen by workers under utility maximization conditions. Rents and population density decrease with distance from the city center. We distinguish between three area types: (i) The central city, closest to the CBD, (ii) the suburb, between the central city and the edge of the city, and (iii) the area outside the city, which is called the periphery, in which population density is exogenously given.

Improvements in the highway network reduce commuting costs per unit distance. Workers then choose a residential location further away from the city center, which reduces population density within the central city (Baum-Snow, 2007b), and increases population in areas which are connected to the highway network.[4]

These effects may change when land policies are introduced. Figure 3.1 illustrates the effects of highways given land development restrictions in designated areas in the outskirts of the city. When no highways or land development restrictions are present, population density decreases in a relatively quick rate with the distance from the city center, until the city edge (t_0). The construction of a new highway will reduce commuting costs and result in outward expansion of the city edge (see t_1), but development restrictions would result in a decline in population density in restricted areas, and population redistribution to adjacent areas (see t_2).

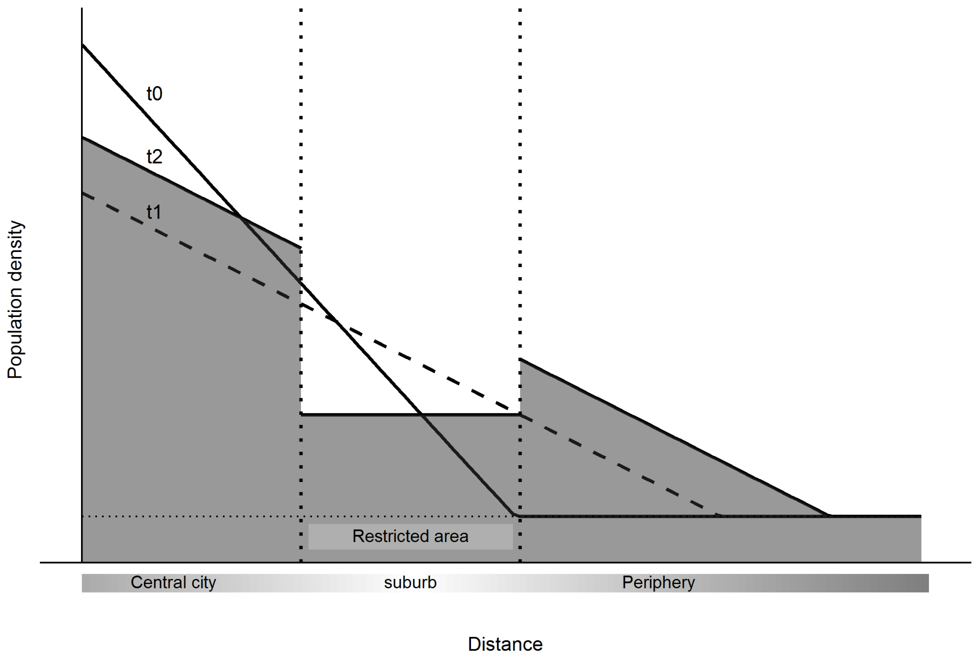
The model predicts that the combined effect on population growth in the central city and suburb areas may be either positive or negative. This depends on the magniude of the positive effect following an increase in highway extent and decrease in commuting costs, as well as on the strictness of the development and its subsequent negative effect on growth. However, for population in peripheral areas the effects of highways and planning restrictions are both

[3]The main consequence is that we assume that total population is not affected by land policies.

[4]Highway improvements may also increase the attractiveness of an urban area, and its ability to attract new workers (Duranton and Turner, 2012).

positive, as population density is expected to increase due to the reduction in commuting costs and the redistribution of population from the restricted area. The combined outcome reflects a leapfrog pattern of low population growth rates in the suburbs, but high growth rates in peripheral areas.

Figure 3.1: Illustration of the theoretical framework



Notes: t_0: No highways, no development restrictions. t_1: Highways, no development restrictions. t_2: Highways and development restrictions.

3.3 Spatial planning and transportation development in the Netherlands

3.3.1 Land development restrictions

The origins of Dutch land use planning date back to the early 20th century when the Housing Law (1901) obliged cities to make a plan for large scale extensions of their residential areas. Areas without such plan could not be developed, at least not on a large scale. However, the planning system was universally implemented after World War II. During the 1950's it became clear to policy makers that the population should be expected to continue to grow in the next decades and that this would have substantial consequences for land use. In line with prevalent ideas about government intervention in the economic system, it was thought that land use planning could contribute to an orderly development of national land use that would increase social welfare.[5]

The 1958 document on 'The Development of the Western Part of the Country' (*Rijksdienst voor het Nationale Plan*, 1958) presents a vision on spatial planning of the Randstad that would remain dominant until the 1980s. "If one allows the current development to proceed, one of the main advantages of the Dutch Randstad in comparison to foreign conurbations will be forfeited: the spatially separately located cities of transparent size". This vision was translated into several main policy measures. First, the center of the urbanized Randstad, the 'Green Heart', was preserved for agricultural use (Koomen et al., 2008; Koomen and Dekkers, 2013). Second, concerns that expansion of large urban areas would result in a formation of one large urban agglomeration were addressed by assignment of 'buffer zones', areas surrounding large cities. Spatial delineation of the zones appeared later in 1966 (Dieleman et al., 1999; Koomen et al., 2008). Third, the central government was also involved in directing urban growth. Areas outside of main cities were defined as growth cities (*'Groeikernen'*), which were destined to absorb suburbanization. The definition of the growth cities and the implementation of the new policy was not fully realized before the late 1970s and early 1980s (following the Third Report on Physical planning, 1974–1977),[6] and was eventually discontinued by the early 1990's when new national policy redirected urban development back to areas in vicinity to traditional city centers (Geurs and Van Wee, 2006; Jobse et al., 1991; Ostendorf and Musterd, 1996).[7]

3.3.2 Urban and road network development

Intercity passenger and freight transportation was historically based on railway and inland waterways. The first railway line was opened in 1839, and the

[5]Spatial planning in the Netherlands is guided by a series of white papers, or Reports on Physical Planning (Nota Ruimtelijke Ordening, released in 1960, 1966, 1974-1977, 1988, and 2001).

[6]The designation of the growth cities (*'groeikernen'*) was soon followed by designation of four growth cities (*'groeisteden'*), existing cities located far from population concentration which were destined to absorb additional urban population growth. We do not make a distinction between groeikernen and groeisteden.

[7]This was directed by the Fourth Report on Physical Planning in the Netherlands (*Vierde Nota Ruimtelijke Ordening*, 1989), and particularly its annex report in 1991 (*Vierde Nota Ruimtelijke Ordening Extra*, abbreviated as 'VINEX').

railway network reached its peak length in 1930 (Koopmans et al., 2012). Motorized transport was quickly adopted after the introduction of automobiles in 1896 (Smaal, 2012). By 1930 there were approximately 68,000 registered private auto-mobiles, 30,000 motorcycles and 44,000 busses and trucks (Veenendaal, 1996). The main roads were then built based on a plan which was first laid out in 1821 (see Appendix 3.A). In 1927 the government laid out its first official road network plan (*'Rijkswegenplan'*).[8] The first highway (between The Hague and Utrecht) was opened in 1937. By 1960 the network reached a length of 351 km (Smaal, 2012). During the same period, between 1946 and 1960, car ownership grew from 5 cars per 1,000 people, to 45.8 cars per 1,000 people (Ligtermoet, 1990), still only a fraction of contemporary car ownership rates in the United States.[9]

Following growth in traffic, the government decided in 1961 to expand the network, and to construct additional 1,200 kilometers of highway before 1975. This goal was reached in 1972 (Ligtermoet, 1990). Car ownership rates also grew rapidly to 218 cars per 1,000 people, approximately half of US car ownership rates at the time. After the completion of the expansion plan, funds for road investments were exhausted. Around the same time, increasing attention to environmental impacts and congestion effects of car travel, and the breaking of the oil crisis in 1973 (after which driving on Sunday was prohibited for two months), led to a new government policy which aimed to decrease the dependency in commuting by car (Ligtermoet, 1990; Schwanen et al., 2004; Smaal, 2012). The expansion of the highway network during the 1960s is now regarded as an outdated policy. The development of the highway network continued at a slower after the 1970s,[10] and it was characterized by the debate between the demand for better roads and environmental preservation and limitation of energy usage.

3.4 Data

We estimate the effect of new highways on population growth distinguishing between (i) central cities (ii) suburban municipalities and (iii) peripheral municipalities. Growth of the highway network occurred primarily between 1961 and 1972. Hence, we will examine population growth just after this period, so between 1970 and 1980. The choice of the exact time period was made to maintain consistency with highway variables, which are only available for ten

[8] This plan was revised several additional times in the following decades, to consider updated projections of traffic and to specify different road capacity types.

[9] By comparison, in 1960 there were 411 cars per 1,000 people in the United States.

[10] See highway network expansion maps in Appendix 3.A.

years intervals (1960, 1970 and 1980). As a sensitivity analysis we also analyze population growth between 1980 and 1990 to examine the long-term effects.

We make use of historical data on the extent of transportation networks and population, calculated for 811 municipalities for 1980 municipal boundaries (using 250 square meter cells). Table 3.1 provides a descriptive summary of the variables used in the analysis.

Our main dependent variable is population growth between 1970–1980 per municipality, which was obtained from Statistics Netherlands. On average, population grew by 22 percent between 1970 and 1980. Data on the 1960, 1970 and 1980 highway network was obtained through the Historisch NWB (*Nationaal Wegenbestand*).[11] Information regarding the main roads in 1821 was available through the ministry of infrastructure and environment (see appendix 3.A). For 1970, road data was used to create two variables that measure highway extent: highway density and rays. The average distance to highway access point was also used as an alternative measure of highway extent. The results of this measure produce similar coefficients, but are less trustworthy, as shown by a low first-stage Kleibergen-Paap F-test score which implies a weak instrument. Highway access points are frequent in the Netherlands, and access to highways is almost continuous. In 1970, the highway network had approximately 340 access points, which corresponds with an access point for every 2.8 highway kilometers on average.

Highway density is calculated as the ratio of the meter length of highways in a municipality and the municipal area (in square kilometer). We calculate highway rays from each municipal area center following Baum-Snow (2007a) and Baum-Snow et al. (2017). We define a 5 kilometers radius around each municipal centroids (defined by neighborhood population weights), and count the number of times highways cross this radius.[12] Table 3.2 provides a descriptive summary of both highway extent variables. Our measures of highway extent complement each other. The use of highway rays to study the effect of highways on suburbanization is common in literature. However, highway density may better reflect highway accessibility in rural areas (particularly if access points are frequently present), or in municipalities with large areas or irregular boundary shapes, in which highways cross the municipal area but do not directly reach population centers.[13] We use both highway extent measures in

[11] Highways were identified based on road type indication of dual motorway or highway, or whether the road is maintained by the central government. Both result in similar figures of highway network length, as reported in Ligtermoet (1990) and Statistics Netherlands (2015).

[12] Radius of 5 kilometers was determined based on common municipality areas. A sensitivity analysis included using rays based on 3 kilometers radius, which have shown little differences in coefficient values and statistical significance of the estimators.

[13] Peripheral municipalities often include several small villages, and therefore can have mul-

Table 3.1: Summary of main variables

Statistic	Year	Mean	St. Dev.	Min	Max
Pop. growth (1970-1980)	1970-1980	0.199	0.254	−0.641	2.927
Pop. growth (1980-1990)	1980-1990	0.068	0.151	−0.5606	1.94
Highway rays	1970	1.321	1.367	0	8
Highway density	1970	68.7	113.5	0	941.7
Rail stations	1930	1.02	1.748	0	18
Population (1930)	1930	9,718.10	38,786.10	0	743,900.7
Population (1960)	1960	14,005.7	48,420.5	34.797	855,539.7
Buffer zone share	1966	0.046	0.182	0	1
Nature coverage share	1960	0.098	0.139	0	0.921
Green heart share	1958	0.148	0.353	0	1
Distance from central city	1980	22.691	15.362	0	95.433
Reclaimed land share	1980	0.014	0.098	0	1
Growth city	1977	0.022	0.147	0	1
Road density (instrument)	1821	54.3	101.8	0	727.4
Artificial buffer zone share (instrument)	1980	0.212	0.409	0	1

Notes: (i) The number of observations is 811. (ii) Highway and road density are defined as length in meters per square km. (iii) Municipalities with zero population in 1930 are in areas which were reclaimed from the sea. (iv) We report population growth (1970–1980) excluding two municipalities, which were built on land reclaimed from the sea. (v) We report population growth (1980–1990) excluding one municipality (Almere), which was built on land reclaimed from the sea.

order to reconcile such possible differences in measurements. We define central cities as cities with a population exceeding 50,000 people in 1930.[14] Since municipalities vary in area, we also restricted the definition of central city to municipalities which had population density level of at least 500 inhabitants per square kilometers in 1930.[15] Suburban municipalities are defined as municipalities directly adjacent to central cities, or located within five kilometers from the centroid of a central city.[16] This definition is in line with the idea that these municipalities are most likely to experience population expansion following improvement in highways. All other municipalities are defined as

tiple town centers.

[14]Appendix 3.B.1 includes a sensitivity analysis with population thresholds of 35,000, and 65,000 inhabitants.

[15]Despite having a population of approximately 60,000 inhabitants in 1930, the municipality of Apeldoorn was not included as a central city. Apeldoorn has the largest municipal territory in the Netherlands with 339 square kilometers of municipal area, of which 81% is open space. With a population density of 176 people per square km in 1930, it is rural in character.

[16]We also tested suburb definitions of municipalities located 10, 15 and 20 kilometers from the centroid of central cities. Highway extent coefficients maintain relatively similar values and their statistical significance level, with the exception of rays under the 20 kilometer radius suburbs scenario, where the effect becomes statistically insignificant.

peripheral municipalities. In total, we define 20 central cities, 133 suburban municipalities and 658 peripheral municipalities (Figure 3.2). The use of historical population levels to define municipality types is also used in order to relief suspicion of endogeneity in the assignment of central cities, suburbs and peripheral municipalities.

Suburban municipalities face development restrictions when they are located within the Green Heart or within buffer zones. Buffer zones cover 12.9 percent of suburban municipalities' area, compared to 7.9 and 2.6 percent of central cities and peripheral municipality's area. The green heart covers 17.4 percent of suburban municipalities' area, compared to 4.6 and 13.8 percent of central cities and peripheral municipalities' area.[17]

We also use information on rail stations as a control variable. Because railway length reached its maximum in 1930 it is possible that the presence of stations in 1930 interfered with the effects of new highways. Data on historical railway stations was obtained from Koopmans et al. (2012). Data on nature coverage in 1960 was available from Alterra (Kramer, 2005). Spatial planning and land development restriction data, including the boundary of the green heart and the buffer zones, was obtained from Koomen et al. (2008).

The Netherlands is characterized by abundance of water, which often obstructed land development. During the 20$^{\text{th}}$ century large land reclamation projects were carried out in the Netherlands, and such lands were often designated for development purposes. To control for population growth in land reclaimed from water we include the share of municipal area which was reclaimed between 1930-1980, which was calculated by comparing water line boundaries between 1930 and 1980.[18]

To guide urban sprawl away from areas designated to remain in natural or agricultural use, the Dutch government defined "growth cities" in which urban development was concentrated. The growth cities lie within eighteen 1980 municipalities, of which two are central cities (Breda, Groningen), 6 are suburban municipalities and 10 are peripheral municipalities. The delineation of the growth cities was implemented in the late 1970's.

[17]The correlation between the share of municipal area included within the buffer zones and the suburb municipalities dummy is 0.209. The correlation between the share of area included within the Green heart and suburb municipalities dummy is 0.355.

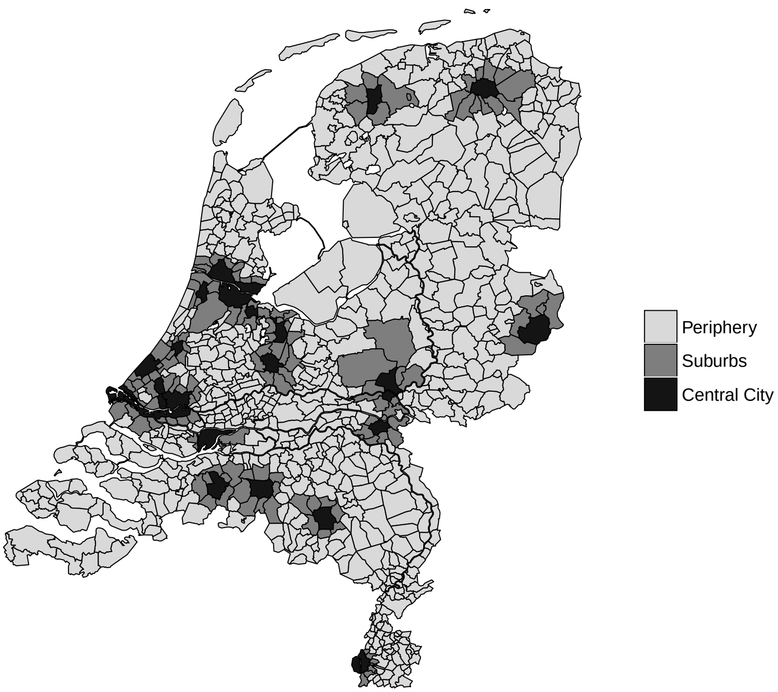
[18]The 1980-1990 analysis includes the same variable, adjusted to 1990 levels.

Table 3.2: Summary of highway variables

	Municipality group	N	Mean	St. Dev.	Min	Max
Highway rays	Central city	20	2.65	1.182	1	5
	Suburb	133	1.789	1.446	0	7
	Periphery	658	1.185	1.315	0	8
Highway density	Central city	20	174.1	98	0	423.1
	Suburb	133	75.6	111.8	0	584.6
	Periphery	658	64.2	112.8	0	941.7

Note: Highway density is defined in meters per square km.

Figure 3.2: Municipalities (1980)



3.5 Methodology

3.5.1 Main specification: Population growth (1970–1980)

We investigate population growth in urban areas by specifying the following specification:

$$\begin{aligned}\Delta \ln(Pop_i) =& \beta_0 + \beta_S S_i + \beta_C C_i + \beta_{HP} H_i P_i + \beta_{HS} H_i S_i + \beta_{HC} H_i C_i \\ &+ \sum_k \alpha_k X_{i,k} + \epsilon_i,\end{aligned} \tag{3.1}$$

where: $\Delta \ln(Pop_i)$ is the change in log population in municipality i. P_i indicates a peripheral municipality, S_i indicates a suburban municipality, and C_i indicates a central city. H_i defines highway extent, highway rays or highway density (in log). H_i is interacted with P_i, S_i and C_i to allow distinct effects of highways by type of municipality. $X_{i,k}$ refers to control variables, including the number of rail stations and population level, both in 1930, nature coverage in 1960, the municipality share within a buffer zone, and the municipality share within the green heart. We will also estimate another specification in which we restrict $\beta_C = \beta_S$ and $\beta_{HC} = \beta_{HS}$. This restricted specification allows us to study the effects of highways in urban agglomerations as a whole, consisting of central cities and suburbs.

We address the concern that the development of the highway network is most likely endogenous, as its construction is likely influenced by travel demand. We use an instrumental approach using the 1821 road network as an instrument (see also Baum-Snow (2007a,b); Baum-Snow et al. (2017); Duranton and Turner (2012); Garcia-López (2012); Garcia-López et al. (2015); Pasidis et al. (2015).[19]

Many spatial development restrictions and land use policies in the 1960s and 1970s are also suspected to be endogenous. This is unlikely the case for buffer zones, because their definition relies on their exogenous location in immediate adjacency to traditional urban centers. Hence, in our main analysis we consider buffer zone share as exogenous. Nevertheless, in a sensitivity analysis we consider them as endogenously determined, employing an (artificial) buffer area of fifteen kilometers around central cities as an instrument. We

[19] The road network system of 1821 was created before the industrial revolution and a century of rail-transport dependency, and thus unlikely to have been affected by the changes in spatial structure that resulted from improvements in transportation technologies in the following century. This argument suggests that rail infrastructure may also be used as a valid instrument for the highway network, as planned highway trajectories often followed existing historic railway lines, particularly where bridge crossing were already present. However, due to its close association with later population growth, it can be argued that rail accessibility does not satisfy the exclusion restriction.

find the results unchanged for the impact of highways. We also take into account that the effect of highways may differ for areas within and outside buffer zones, and use additional interaction terms between highways, buffer zones and municipality types.

Our sample includes eighteen growth cities that were designated late during our study period (1977), and are likely to be dependent on the development of highway accessibility at the time. We therefore do not control for the presence of growth cities. Our results are also robust when these eighteen municipalities are excluded from the analysis.

3.5.2 Endogeneity issues of the interaction terms

The above specification includes interactions between highways extent variables and three municipality types: central cities, suburbs and peripheral municipalities. It follows that the three interaction terms H_iP_i, H_iS_i, H_iC_i may be considered as endogenous. In the context of an endogenous variable with interactions, several estimation procedures are then possible.

The standard approach, which we will refer to as the multiple first-stage approach, is to estimate a separate first-stage regression for each of the endogenous interaction variables.[20] A second approach, which we will show to be more efficient, is to estimate a single first-stage to predict $\hat{H}_i$ using an instrument, denoted by Z_i, and to use this predicted variable interacted with the dummies for central cities, suburbs and peripheral municipalities in the second stage (as suggested by Balli and Sorenson, 2013). Hence, one estimates the following first-stage equation:

$$H_i = \delta_0 + \delta_S S_i + \delta_C C_i + \delta_Z Z_i + \sum_k \gamma_k X_{i,k} + u_i, \tag{3.2}$$

where δ and γ_k are coefficients to be estimated. Given the first-stage estimates of the coefficients in (3.2), $\hat{H}_i$ is predicted, and is then interacted with P_i, S_i and C_i respectively. The resulting interaction terms $(\widehat{H_iP_i}, \widehat{H_iS_i}, \widehat{H_iC_i})$ are then used in a second stage. Robust standard errors in the second stage can be calculated following Angrist and Pischke (2008).

Because P_i, S_i, and C_i are mutually exclusive categorical variables (i.e, dummy variables that sum to one), the approach generates consistent and efficient estimators. To demonstrate this, first consider a multiple first-stage approach in which all coefficients are allowed to vary for periphery, suburbs

[20]See Wooldridge (2002), page 122.

or central cities:

$$P_iH_i = \quad \delta_P P_i + \delta_{P,Z} P_i Z_i + \textstyle\sum_k \gamma_{P,k} P_i X_{i,k} + u_{p,i} \tag{3.3a}$$
$$S_iH_i = \quad \delta_S S_i + \delta_{S,Z} S_i Z_i + \textstyle\sum_k \gamma_{S,k} S_i X_{i,k} + u_{s,i} \tag{3.3b}$$
$$C_iH_i = \quad \delta_C C_i + \delta_{C,Z} C_i Z_i + \textstyle\sum_k \gamma_{C,k} C_i X_{i,k} + u_{c,i} \tag{3.3c}$$

where $P_i + S_i + C_i = 1$. It is well known that using (3.1) and (3.3) together generates consistent estimators (Wooldridge, 2002, p. 122). These first stages can equivalently be estimated jointly by summing equations (3.3a), (3.3b) and (3.3c):

$$\begin{aligned} H_i = \delta_P P_i + \delta_S S_i + \delta_C C_i + (\delta_{P,Z} P_i + \delta_{S,Z} S_i + \delta_{C,Z} C_i) Z_i \\ + \sum_k (\gamma_{P,k} P_i + \gamma_{S,k} S_i + \gamma_{C,k} C_i) X_{i,k} + u_i, \end{aligned} \tag{3.4}$$

where $u_i = u_{p,i} + u_{s,i} + u_{c,i}$. Hence, the single first-stage approach based on (3.2) is a special case of the multiple-first stage approach, given the restriction $\delta_Z = \delta_{P,Z} P_i + \delta_{S,Z} S_i + \delta_{C,Z} C_i$ and $\gamma_k = \gamma_{P,k} P_i + \gamma_{S,k} S_i + \gamma_{C,k} C_i$ for all $k = 1 \ldots K$. This restriction can be tested using a standard F-test. If it holds, the single first-stage approach is more efficient than the multiple first-stage approach, as follows from a general econometric argument that imposing valid restrictions improves the efficiency of estimators. Furthermore, note that the single first-stage approach is less likely to suffer from weak instruments (as it avoids using additional instruments for each endogenous interaction term). Therefore there may be cases where the approach is not only more efficient, but it is also more likely to produce less biased results.

While both the multiple and single-first stage approaches assume that endogeneity in H_i results in endogeneity of the interaction terms, we also examine a third approach which treats the interaction terms as exogenous, under certain conditions. Bun and Harrison (2014) show that an interaction term between endogenous and exogenous variables may be treated as exogenous if the conditional expectation of the endogenous variable and the error term ($H_i u_i$) does not depend on the exogenous regressor with which the endogenous variable is interacted (P_i, S_i, C_i). The validity of this assumption can be tested using a standard Hausman test (Hausman, 1978), or an extended version which also tests the presence of weak instruments (Bun and Harrison, 2014; Hahn et al., 2011). This test compares the results of the model where the interaction term is treated as endogenous with the model in which it is treated as exogenous.[21]

[21] Following Bun and Harrison (2014), we also use Wald test to test the hypothesis of consistency of OLS estimators, in which all variables are considered exogenous. This hypothesis is rejected.

We will apply all three approaches. It appears that these approaches result in similar coefficient values. The single first-stage approach produces the smallest standard errors and the highest first stage Kleibergen-Paap F-test value, and is therefore preferred.

3.5.3 Peripheral municipalities subsample

The literature on the effects of highways in peripheral regions commonly treats the assignment of highways as exogenous, as it is argued that highway assignment depends on the exogenous location between two larger population centers that are connected with a highway.[22] Plausibly, this exogeneity assumption does not hold in our analysis. Due to the relatively small spatial scale, it is likely that highway assignment in peripheral areas is endogenous. Trajectories of planned highways may have been directed to pass through faster growing peripheral towns, where the spatial scale is sufficiently small, such network planning decisions can be done without imposing costly road bypasses. To study the effects of highways in peripheral municipalities we estimate (3.1) on a restricted subsample of the peripheral municipalities. We will instrument highway extent as in (3.4).

3.6 Estimation results

3.6.1 Main results: Population growth (1970–1980)

To apply the single first-stage approach, we first test whether the restrictions imposed are valid using a standard F-test. The results of the hypothesis test of these restrictions show that for both highway measures we cannot reject the null hypothesis, which indicates that the single first stage approach is valid.[23]

The estimation results of model (3.1) using the single first-stage approach, are presented in Table 3.3.[24] The results in columns 1 and 2 show that one highway ray increases municipal population growth by 15.9%. One percentage increase in highway density is expected to increase population growth by 0.058%. This corresponds with an increase of approximately 10% given one standard deviation increase in highway density. The effect of highway rays and

[22]See, for example, Chandra and Thompson (2000); Fajgelbaum and Redding (2014); Michaels (2008).

[23]Both F-tests show values lower than 0.8, so with a corresponding p-value exceeding 0.7. The F-test has 17 degrees of freedom. We restrict the nine coefficients of highway extent (δ_Z) and eight control variables (γ_k) to be equal between municipality types. Because distance to central cities has no variation within central cities, we have 17 restriction.

[24]First stage estimation results show a positive and statistically significant effect of the instrument on highway extent in the 1970s, see Table 3.C.1.

density on population growth in central cities and suburbs is small and statistically insignificant. The effects of buffer zone are negative and significant. An increase in 1% in share of municipal area defined as a buffer zone is expected to drop population growth by approximately 12.7-14.9%. This confirms the findings of Geurs and Van Wee (2006); Koomen et al. (2008); Koomen and Dekkers (2013) that the Dutch policy of open space preservation was effective in preventing urban sprawl in designated areas. A positive coefficient for suburbs and central cities (column 2) indicates that urban municipalities experienced stronger population growth compared with peripheral areas.

The results of the restricted specification (columns 3 and 4), in which we assume an identical effect of highways for agglomerated urban areas (central cities and suburbs), also show positive effects of highways on peripheral population growth. The effect of increase in highway rays on population growth is estimated at 17.2%. The effect of one percentage increase in highway density is essentially also the same, at 0.064%. We find a significant effect of highway density on population in urban agglomerations (central cities and suburbs, combined), but this effect disappears when highway rays are examined.

The result that highways have hardly any effect on central city population, but a strong effect on the peripheral municipalities, largely confirms the findings of Baum-Snow (2007a,b) and related papers. However, the absence of a clear effect on population growth in suburban municipalities is not in line with previous findings. A possible explanation is the small spatial scale in the Netherlands. It is possible that municipalities adjacent to central cities do not experience a significant reduction in commuting costs following the construction of a new highway, and commuters from these municipalities still prefer local urban roads, or public urban transportation systems. Moreover, it may be that commuting costs remain relatively low in municipalities located further away from central cities, in that sense, the strong positive effect of highways found in peripheral municipalities could be interpreted as reflecting this suburbanization effect.[25]

We estimate an additional specification of the model in which we consider buffer zone share to be endogenous. Here we instrument buffer zones using an 'artificial buffer' variable, defined by a dummy which indicates municipalities that are completely contained within a radius of 15km from the cities around which buffer zones were present (Amsterdam, Rotterdam, The Hague, Utrecht and Maastricht).[26] The estimation results (Table 3.3, columns 5-6) show that

[25] Twelve suburban municipalities are adjacent to two central cities. We have also considered a specification in which these suburban municipalities experience a double effect. The results maintain very similar values, and as expected, the estimated effect of H_iS_i becomes slightly smaller.

[26] We also experimented with artificial buffers in ranges between 3-20 kilometers radius from

Table 3.3: Population growth and highways – Main results

	Dependent variable: Log. pop. growth (1970–1980)					
	Main model		Restricted specification		Main model	
	(1)	(2)	(3)	(4)	(5)	(6)
Highway rays*periphery	0.15982***		0.17245***		0.15262***	
	(0.06184)		(0.05843)		(0.05596)	
Highway rays*suburb	−0.00456		−0.03959		−0.00780	
	(0.07982)		(0.06623)		(0.06701)	
Highway rays*central city	−0.01132		−0.03959		−0.08769	
	(0.10705)		(0.06623)		(0.11379)	
Highway density*periphery		0.05877***		0.06461***		0.05638**
		(0.02155)		(0.02050)		(0.02393)
Highway density*suburb		−0.02957		−0.04516*		−0.01423
		(0.02980)		(0.02396)		(0.02266)
Highway density*central city		−0.03991		−0.04516*		−0.04079
		(0.02919)		(0.02396)		(0.03322)
Suburb	0.28349	0.21933*	0.37373***	0.29253***	0.29345**	0.17887**
	(0.18291)	(0.11396)	(0.13702)	(0.09324)	(0.13192)	(0.07812)
Central city	0.16184	0.11401	0.37373***	0.29253***	0.34126	0.11201
	(0.26903)	(0.14813)	(0.13702)	(0.09324)	(0.29880)	(0.16357)
Buffer zone share	−0.14900***	−0.12744***	−0.13571**	−0.12089***	−0.38633**	−0.18645
	(0.05225)	(0.03978)	(0.05264)	(0.04071)	(0.17001)	(0.13901)
Rail stations	0.00238	0.00152	0.00110	0.00006	0.00119	0.00095
	(0.00345)	(0.00354)	(0.00357)	(0.00369)	(0.00338)	(0.00362)
Log Pop. (1930)	0.00519	−0.00658	0.00374	−0.00938	−0.00123	−0.00790
	(0.03940)	(0.03624)	(0.03965)	(0.03604)	(0.03972)	(0.03855)
Log Pop. (1960)	−0.05521	−0.04869	−0.05821	−0.05421	−0.05175	−0.04844
	(0.04862)	(0.04638)	(0.04798)	(0.04599)	(0.04909)	(0.05004)
Nature coverage share	−0.06932	−0.04736	−0.06344	−0.03564	−0.09461**	−0.05597
	(0.04530)	(0.05008)	(0.04465)	(0.04996)	(0.04719)	(0.05173)
Green heart share	−0.03986	0.02458	−0.03631	0.03189	−0.04069	0.02200
	(0.03806)	(0.02279)	(0.03880)	(0.02351)	(0.03893)	(0.02386)
Distance from central city	−0.00125	−0.05355***	0.01677	−0.03071**	−0.01464	−0.05682***
	(0.01876)	(0.01374)	(0.01560)	(0.01258)	(0.01781)	(0.01465)
Reclaimed land share	0.60556*	0.52545	0.58282*	0.49228	0.57074	0.52261
	(0.34107)	(0.32010)	(0.34253)	(0.31686)	(0.34757)	(0.33426)
Constant	0.41606***	0.68340***	0.38200***	0.66955***	0.49993***	0.71033***
	(0.09289)	(0.13068)	(0.09197)	(0.13131)	(0.10734)	(0.13993)
Highways instrumented	Yes	Yes	Yes	Yes	Yes	Yes
Buffer zone instrumented	No	No	No	No	Yes	Yes
Kleibergen-Paap statistic (Highways)	10.96	33.62	10.97	32.50	7.994	13.75
Kleibergen-Paap statistic (Buffer zone)					16.15	16.15
Observations	811	811	811	811	811	811

Notes: (i) Highway extent variables are in 1970 levels. (ii) Highway density is expressed in logarithm, (iii) Possible endogeneity in the interaction terms is addressed following the single first-stage approach (Balli and Sorensen, 2013). (iv) Robust standard errors in parentheses. $p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$

the effect of highways in peripheral municipalities remains robust when we consider buffer zones to be endogenous.[27]

3.6.2 Endogeneity in the highways interaction terms

In Table 3.4 we present the results of other approaches to address endogeneity in the interaction terms. We compare the results of the main model as presented in Table 3.3 using (i) OLS, (ii) assuming that the interaction terms are exogenous (following the approach of Bun and Harrison, 2014), and (iii) assuming endogenous interaction terms, and estimating separate first steps for each term as commonly applied (e.g, Wooldridge, 2002). Note that the Kleibergen-Paap F-test of the single first-stage approach (presented in Table 3.3) is higher than the values of the test for exogenous interaction terms and multiple first-stage approaches.

The results show that for all instrumented approaches, and for both highway measures, the estimated interaction coefficient with peripheral municipalities obtains similar values. Note that the multiple first-stage approach for highway density (columns 3 and 6) are invalid, as demonstrated by a low F-test value, implying that the results based on this approach are biased.

Notably, the effects of highways are found to be substantially lower in value when estimated in OLS (columns 1 and 4) compared with the instrumented estimations. Since reversed causality is controlled, and would be expected to result in overestimation rather than underestimation of the effects, the lower coefficient value is possibly related to omitted variable bias. For instance, it may be that highway planners initially decided that highway trajectory would reach 'growing' population centers, but without entering them, in order to allow highway access while avoiding expected negative highway noise externalities.

3.6.3 Peripheral municipalities subsample

The results using a subsample of peripheral municipalities are presented in the 3.B.2, Table 3.B.3. They show similar effects as in Table 3.3. Since the previous literature generally considers highways in peripheral areas as exogenous, we also estimate the restricted specification using OLS. The results show that the effects of highways become substantially smaller. The effect of an increase in one highway ray is about 1.38%, and the effect of 1% increase

central cities. A buffer of fifteen kilometers was chosen as it is in line with actual buffer zones ranges.

[27] A standard Hausman test for the endogeneity of buffer zone share provides a value of 1.91, corresponding with a p-value of 0.166 which suggests that the exogeneity assumption is not rejected.

Table 3.4: Analysis of endogeneity of interaction terms

Dependent variable: Log. pop. growth (1970-1980)

	OLS	Exogenous interactions	Multiple first-stage	OLS	Exogenous interactions	Multiple first-stage
	(1)	(2)	(3)	(4)	(5)	(6)
Highway rays*periphery	0.01593***	0.12191**	0.12674**			
	(0.00435)	(0.05429)	(0.05952)			
Highway rays*suburb	0.00384	0.00868	0.12491			
	(0.01678)	(0.01738)	(0.14907)			
Highway rays*central city	−0.01301	−0.01122	−0.31745			
	(0.01383)	(0.01430)	(1.07391)			
Highway density*periphery				0.01228***	0.04115***	0.04565
				(0.00288)	(0.01535)	(0.11888)
Highway density*suburb				−0.00163	0.00012	0.04594
				(0.00897)	(0.00878)	(0.45221)
Highway density*central city				−0.01271*	−0.01296**	−5.49310
				(0.00676)	(0.00656)	(275.74290)
Suburbs	0.03703	0.18742**	0.00169	0.04907	0.10486*	−0.00718
	(0.02668)	(0.08289)	(0.21536)	(0.04139)	(0.05693)	(1.39078)
Central city	−0.18358***	0.02790	0.87568	−0.15349**	−0.06437	26.40456
	(0.06358)	(0.14063)	(2.71406)	(0.06341)	(0.08639)	(1,330.757)
Buffer zone share	−0.12058***	−0.13671***	−0.17805*	−0.11564***	−0.11559***	−0.00568
	(0.04012)	(0.04573)	(0.09083)	(0.03738)	(0.03908)	(6.46219)
Rail stations (1930)	0.00354	0.00382	0.00824	0.00305	0.00197	0.04551
	(0.00335)	(0.00461)	(0.01401)	(0.00338)	(0.00353)	(2.25659)
Log Pop. (1930)	−0.02194	−0.00029	0.00682	−0.02187	−0.01333	0.00246
	(0.03384)	(0.04037)	(0.04174)	(0.03286)	(0.03404)	(0.66244)
Log Pop. (1960)	−0.01436	−0.04393	−0.05765	−0.01758	−0.03712	−0.09695
	(0.03890)	(0.04755)	(0.05293)	(0.03719)	(0.04145)	(2.62362)
Nature coverage share	−0.06825	−0.05248	−0.07327	−0.06229	−0.04613	−0.14641
	(0.04453)	(0.05356)	(0.06851)	(0.04521)	(0.04967)	(4.26074)
Green heart	0.02011	−0.03407	−0.03946	0.02910	0.02933	0.01873
	(0.02296)	(0.04338)	(0.05314)	(0.02228)	(0.02308)	(0.03996)
Distance from central city	−0.04276***	−0.02120	−0.01385	−0.04780***	−0.05165***	−0.05590
	(0.01294)	(0.01983)	(0.02673)	(0.01319)	(0.01465)	(0.19903)
Reclaimed land share	0.48798	0.60396*	0.64925*	0.47617	0.49514	0.60541
	(0.31882)	(0.34187)	(0.34611)	(0.31352)	(0.30973)	(4.57343)
Constant	0.58241***	0.46841***	0.49801***	0.61804***	0.66816***	1.02210
	(0.11289)	(0.11834)	(0.13249)	(0.10879)	(0.12072)	(15.54208)
Instrumented	No	Yes	Yes	No	Yes	Yes
Kleibergen-Paap statistic		3.618	0.037		9.005	0.0001
Observations	811	811	811	811	811	811
R^2	0.1691			0.1755		

Notes: (i) Instrumented variables in columns 2 and 6: highways, interacted in the second stage with regional dummies. Instrumented variables in columns 3 and 7: highway*periphery. Instrumented variables in columns 4 and 8: highway*periphery, highway*suburbs, highway*central city. (ii) Highway extent variables are in 1970 levels. (iii) Highway density is in logarithm. (iv) Robust standard errors in parentheses. *$p<0.1$, **$p<0.05$, ***$p<0.01$

in highway density is about 0.0105%, indicating a negative bias. This implies that in contrast with previous assumptions in other studies, highway extent in peripheral areas in the Netherlands cannot be regarded as exogenous.[28]

3.7 Sensitivity analysis and long-term effects

3.7.1 Interactions with buffer zones

Our main model specification tests the effects of highways for several municipality types, controlling for presence of buffer zones. It may also be that the effects of highways vary with the presence of buffer zones. To test this we extend our model by including an additional set of interaction variables between highway extent and buffer zones.[29] First, we assume that buffer zone share is exogenous. We later test our results under an endogenous buffer zone share assumption. As before, our instrumental variable strategy follows the single first-stage approach to compute each of the three instruments (for the three endogenous variables - highway extent, buffer zone share and highways$*$buffer zone share). The instruments are first estimated using first-stage regressions, and then interacted in the second-stage with the three municipality type dummies.

The results in Table 3.5 (columns 1 and 2) show positive, but weakly significant, effects of highways in buffer zones within peripheral and suburban municipalities. Highway density has a positive effect of 0.15 within suburban buffer zones, significant at the 10% level. The effect of highways in buffer zones within central cities is positive (and statistically significant). Note that this effect is estimated based on only 11 observations with non-zero buffer zone share, so is not reliable. Notably, the effect of highways in peripheral areas remains robust and maintains similar values (and significance levels) when highway and buffer zones interactions are included.

When the buffer zone share is considered endogenous (Table 3.5, columns 3 and 4), the effects of highway density in peripheral buffer zone becomes negative and statistically significant, with a value of -0.20. This implies that due to the presence of buffer zones, the positive effect of highways in peripheral areas is nullified.[30] This means that highways reduced population growth

[28]Excluding eighteen observations that refer to growth cities generates almost identical results.

[29]We focus here on buffer zone share as it is found to have a strong negative effect on population growth. The other type of regulation, the green heart, did not have any effect on population growth and is therefore a less interesting variable to explore.

[30]The combined effect of highways in peripheral buffer zones is calculated as $0.0705 - 0.2078 = -0.137$. However, a standard F-test shows that we cannot reject the null hypothesis that

Table 3.5: Population growth and highways - Effect of highways in restricted areas

(Dependent variable: Log. pop. growth (1970–1980))

	Main model with interaction highways*buffer zones			
	Endogenous highways		Endogenous highways and buffer zone share	
	(1)	(2)	(3)	(4)
Highway rays*periphery	0.16111**		0.14408***	
	(0.06555)		(0.04438)	
Highway rays*suburb	−0.03784		0.00883	
	(0.09862)		(0.07580)	
Highway rays*central city	−0.15251		0.01536	
	(0.11119)		(0.15035)	
Highway rays*periphery*buffer	0.09540		0.43895	
	(0.13122)		(0.84960)	
Highway rays*suburb*buffer	0.29416		−0.33480	
	(0.19135)		(0.48195)	
Highway rays*central city*buffer	1.44828**		0.66940	
	(0.64644)		(1.32445)	
Highway density*periphery		0.05595***		0.07059***
		(0.02128)		(0.01917)
Highway density*suburb		−0.04437		−0.10010*
		(0.03264)		(0.05670)
Highway density*central city		−0.09755***		−0.00037
		(0.02281)		(0.02954)
Highway density*periphery*buffer		0.06949		−0.20789**
		(0.06575)		(0.10364)
Highway density*suburb*buffer		0.15013*		0.46608
		(0.08111)		(0.29378)
Highway density*central city*buffer		0.61490***		0.37623
		(0.16426)		(0.99265)
Suburbs	0.34050	0.25702**	0.24772**	0.41212***
	(0.22361)	(0.12100)	(0.11825)	(0.14166)
Central city	0.52024*	0.36907***	−0.10619	−0.32836
	(0.28313)	(0.11664)	(0.40012)	(0.42583)
Periphery*buffer	−0.28910	−0.16499	−1.31957	0.34247
	(0.20624)	(0.10998)	(1.75975)	(0.27970)
Suburb*buffer	−0.80361*	−0.59905**	0.47317	−1.50225**
	(0.41060)	(0.24148)	(1.04017)	(0.63132)
Central city*buffer	−3.95784**	−2.97854***	0.24209	1.15492
	(1.81640)	(0.80984)	(2.44797)	(0.85749)
Constant	0.42278***	0.67607***	0.59248***	0.62789***
	(0.09445)	(0.13112)	(0.21033)	(0.14379)
Highways instrumented	Yes	Yes	Yes	Yes
Buffer zone instrumented	No	No	Yes	Yes
Highways * buffer zone instrumented	No	No	Yes	Yes
Kleibergen-Paap statistic (Highways)	11.05	33.86	7.994	13.75
Kleibergen-Paap statistic (buffer)			16.15	16.15
Kleibergen-Paap statistic (Highways*buffer)			11.06	7.423
Observations	811	811	811	811

Notes: (i) Included control variables are identical to the variables included in Tables 3.4 and 3.B.3. (ii) Highway extent variables are in 1970 levels. (iii) Highway density is expressed in logarithm. (iv) Robust standard errors in parentheses (v) $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

in peripheral buffer zones, compared with peripheral municipalities outside buffer zones, with similar highway density. The effect of the interactions of highway rays with buffer zone share in all municipality types is statistically insignificant, suggesting that all the effects are captured by the main effects. Furthermore, the effect of highways in peripheral municipalities remains robust. The effect of highway density in suburban municipalities is negative at the 10% level, weakly suggesting that highways have contributed to slower population growth rates in suburban areas, consistent with a leap-frog pattern.

3.7.2 Effects on population growth (developments after 1980)

We have also estimated our model to examine the effects of highways on population growth one decade later, between 1980-1990 (Table 3.B.4). It appears that there is no statistically significant effect of highway extent in the full sample, whereas the effects become weakly significant when we use the subsample of peripheral municipalities.[31]

Possible reasons for the absence of evidence for strong highways effects one decade later is that much of the redistribution of population might have already taken place before 1980, during the years immediately following the great expansion of the highway network. This would suggest that highway effects estimated on the 1970–1980 data represent a long-run effect of approximately 10 to 20 years, after which the effect on population redistribution stabilizes and a new equilibrium is reached.[32]

Our interpretation is that Dutch planning policies aimed to prevent the formation of a large urban conurbation and to preserve agricultural activities and nature. While this policy was effective in achieving its original objectives, it had additional consequences when the highway network was extended. The restrictions that were imposed were compensated by strong population growth in peripheral areas, which resulted in increased commuting distances and time (Schwanen et al., 2001, 2004). Spatial policies in the following decades addressed this by enforcing stricter development restrictions regarding the preservation of open space, and attempting to direct urban population and employment growth back to existing urban areas.[33]

this combined effect is zero.

[31] The effects of buffer zones and nature coverage remain negative and significant, indicating that the presence of development restrictions continues to determine population growth in this period as well.

[32] An additional explanation is that the period 1980-1990 was characterized by national policies which aimed to reduce private car dependency and promote awareness of road externalities. This is also reflected in deceleration in car ownership compared with previous decades (see Figure 3.A.3).

[33] See discussion in Dieleman et al. (1999); Geurs and Van Wee (2006).

3.8 Conclusion

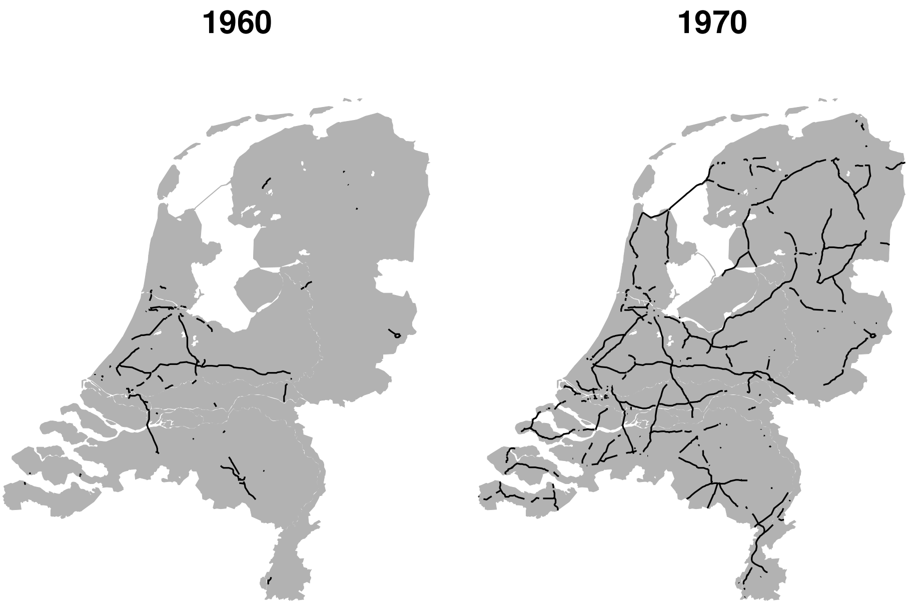
There is a large literature which shows that new highways depopulate city centers. We examine this in the context of the Netherlands where land development restrictions are common. Our analysis focuses on the expansion of the highway network in the 1960's, and its effects on population growth in central cities, suburbs and peripheral areas. We have addressed endogeneity issues by using 1821 road data as an instrument, and employed several innovative approaches to deal with endogeneity in the interaction of highway measures and different types of municipalities (central cities, suburbs and peripheral). In contrast to the literature, our findings for the Netherlands show no effect of highways for central cities and suburban areas. This finding is in line with the idea that strict planning policies and land development restrictions strongly interfere with the effects of highways. In line with the literature, our results show that the rapid development of the Dutch highway network had a substantial effect on changes in the population distribution, and that highways accelerated population growth in peripheral areas by about 10-15 percent. Hence, our results imply that when land development is restricted in the surroundings of cities, new highways divert population growth to locations further away from central cities. The development of the highway network results then in a large scale sprawl, which skips the suburbs and 'leapfrogged' to peripheral towns.

3.A Road networks

Figure 3.A.1: Main roads in the Netherlands (1821)

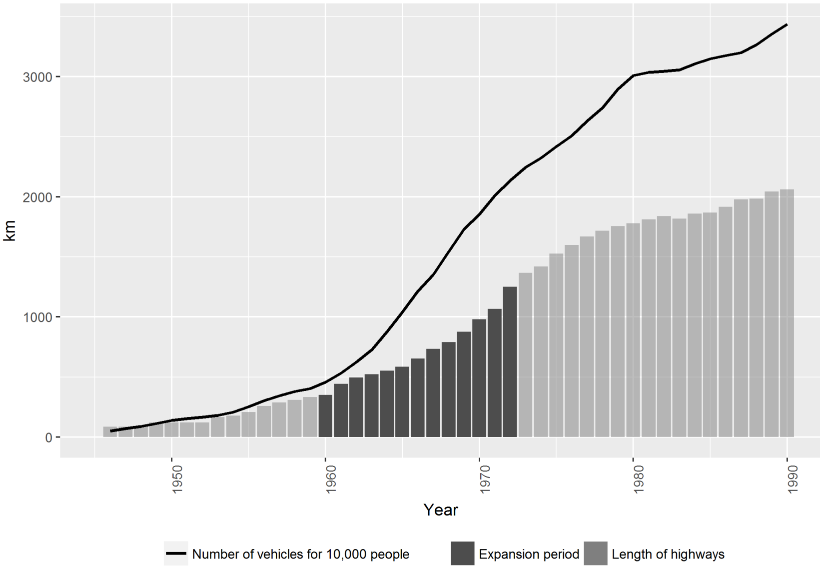
Source: Rijkswaterstaat Beeldbank, multimedia archive (2015) - *Kaart der grote wegen van de 1e klaste met zij kalkvers volgens W.B. 13 maart 1821.* [https://beeldbank.rws.nl/MediaObject/Details/344244 , accessed: December 18th 2016]

Figure 3.A.2: The highway network in 1960 and 1970



Source: Nationaal Historisch Wegenbestand , Data portal of the Dutch government (2015) [https://data.overheid.nl/data/dataset/nationaal-historisch-wegenbestand, accessed: December 19th 2016]

Figure 3.A.3: Highway network length and car ownership (1945–1990)



Source: Ligtermoet (1990), Statistics Netherlands (2016)

3.B Sensitivity analysis

3.B.1 Definition of central cities and suburbs

Table 3.B.1: Comparison between different population thresholds for central cities in 1930

Population threshold (1930)	Central cities	Suburban municipalities	Peripheral municipalities
35,000	26	156	629
50,000	20	133	658
65,000	15	116	680

Table 3.B.2: Population growth and highways – Main results under various central cities population thresholds

Dependent variable: Log. pop. growth (1970–1980)

	Pop. Threshold 35,000		Pop. Threshold 50,000		Pop. Threshold 65,000	
	(1)	(2)	(3)	(4)	(5)	(6)
Highway rays*periphery	0.16027***		0.15982***		0.15090**	
	(0.06123)		(0.06184)		(0.05875)	
Highway rays*suburbs	0.02285		−0.00456		0.03348	
	(0.06971)		(0.07982)		(0.07783)	
Highway rays*central city	0.01199		−0.01132		0.02590	
	(0.07360)		(0.10705)		(0.12592)	
Highway density*periphery		0.05676***		0.05877***		0.05553***
		(0.02109)		(0.02155)		(0.02079)
Highway density*suburb		−0.00935		−0.02957		−0.02026
		(0.02374)		(0.02980)		(0.03059)
Highway density*central city		−0.01483		−0.03991		−0.05496*
		(0.02129)		(0.02919)		(0.02946)
Suburbs	0.22993	0.16774*	0.28349	0.21933*	0.22409	0.20796*
	(0.15180)	(0.09290)	(0.18291)	(0.11396)	(0.18224)	(0.11959)
Central city	0.10100	0.00906	0.16184	0.11401	0.08231	0.21315*
	(0.16243)	(0.09848)	(0.26903)	(0.14813)	(0.30149)	(0.12404)
Buffer zone share	−0.15650***	−0.13027***	−0.14900***	−0.12744***	−0.17187***	−0.12693***
	(0.05353)	(0.04044)	(0.05225)	(0.03978)	(0.05241)	(0.03841)
Rail stations (1930)	0.00254	0.00111	0.00238	0.00152	0.00232	0.00070
	(0.00335)	(0.00351)	(0.00345)	(0.00354)	(0.00343)	(0.00357)
Log Pop. (1930)	0.00500	−0.00732	0.00519	−0.00658	0.00423	−0.00826
	(0.03934)	(0.03627)	(0.03940)	(0.03624)	(0.03934)	(0.03620)
Log Pop. (1960)	−0.05486	−0.04862	−0.05521	−0.04869	−0.06010	−0.05089
	(0.04926)	(0.04732)	(0.04862)	(0.04638)	(0.04895)	(0.04616)
Nature coverage share	−0.06976	−0.05071	−0.06932	−0.04736	−0.07692*	−0.04937
	(0.04480)	(0.04847)	(0.04530)	(0.05008)	(0.04501)	(0.04986)
Green heart share	−0.04220	0.03208	−0.03986	0.02458	−0.04325	0.02610
	(0.04012)	(0.02247)	(0.03806)	(0.02279)	(0.03806)	(0.02248)
Distance from central city	−0.00158	−0.04670***	−0.00125	−0.05355***	0.00365	−0.04365***
	(0.01595)	(0.01344)	(0.01876)	(0.01374)	(0.01633)	(0.01409)
Reclaimed land share	0.60774*	0.52549	0.60556*	0.52545	0.59523*	0.51264
	(0.34038)	(0.32002)	(0.34107)	(0.32010)	(0.33949)	(0.31943)
Constant	0.41488***	0.66496***	0.41606***	0.68340***	0.45961***	0.69286***
	(0.09461)	(0.13886)	(0.09289)	(0.13068)	(0.08792)	(0.13155)
Instrumented	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap statistic	11.54	37.47	10.96	33.62	12.25	34.65
Observations	811	811	811	811	811	811

Notes: (i) Highway variables are in 1970 levels. (ii) Highway density is in logarithm. (iii) Possible endogeneity in the interaction terms is addressed following the single first-stage approach. (iv) Robust standard errors in parentheses. $^{*}p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$

3.B.2 Peripheral municipalities subsample

Table 3.B.3: Population growth and highways – Peripheral municipalities subsample

(Dependent variable: Log. pop. growth (1970–1980)

	IV (1)	IV (2)	OLS (3)	OLS (4)
Highway rays	0.08237**		0.01385***	
	(0.03812)		(0.00390)	
Highway density		0.02868**		0.01052***
		(0.01327)		(0.00271)
Buffer zone share	−0.10029	−0.05930	−0.06766	−0.06042
	(0.06120)	(0.05397)	(0.05381)	(0.05408)
Rail stations	0.00858**	0.00597	0.00812*	0.00727*
	(0.00421)	(0.00419)	(0.00427)	(0.00414)
Log Pop. (1930)	−0.02938	−0.04091	−0.04660*	−0.04672*
	(0.02905)	(0.02717)	(0.02676)	(0.02606)
Log Pop. (1960)	−0.00113	0.00612	0.02331	0.02013
	(0.03358)	(0.03178)	(0.02847)	(0.02795)
Nature coverage share	−0.00194	−0.00084	−0.01462	−0.01119
	(0.05129)	(0.05145)	(0.04924)	(0.04894)
Green heart share	−0.01882	0.03551	0.02775	0.03656
	(0.03720)	(0.02395)	(0.02350)	(0.02353)
Distance from central city	−0.03638***	−0.05444***	−0.04820***	−0.05200***
	(0.01340)	(0.01315)	(0.01278)	(0.01286)
Reclaimed land share	0.37810	0.27595	0.27484	0.26202
	(0.29637)	(0.28204)	(0.27953)	(0.27658)
Constant	0.42688***	0.55136***	0.47355***	0.50807***
	(0.09082)	(0.10886)	(0.09552)	(0.09414)
Instrumented	Yes	Yes	No	No
Kleibergen-Paap statistic	10.04	23.79		
Observations	658	658	658	658
R^2			0.1298	0.1365

Notes: (i) Highway extent variables are in 1970 levels. (ii) Highway density is expressed in logarithm. (iii) Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01

3.B.3 Effects on population growth after 1980

Table 3.B.4: Population growth and highways (1980–1990)

Dependent variable: Log. pop. growth (1980–1990)				
	IV (Full sample)		IV (Periphery subsample)	
	(1)	(2)	(3)	(4)
Highway rays*periphery	0.01473		0.04492	
	(0.02163)		(0.02806)	
Highway rays*suburbs	0.04474			
	(0.06457)			
Highway rays*central city	0.08919			
	(0.12148)			
Highway density*periphery		0.01106		0.01958*
		(0.00987)		(0.01187)
Highway density*suburb		0.01255		
		(0.02638)		
Highway density*central city		−0.00480		
		(0.01965)		
Suburbs	−0.08838	−0.02190		
	(0.14764)	(0.08198)		
Central city	−0.35291	−0.02221		
	(0.40563)	(0.10662)		
Buffer zone share	−0.13296**	−0.11177**	−0.28818**	−0.22665**
	(0.06338)	(0.04480)	(0.12102)	(0.08912)
Nature coverage share	−0.02226	−0.02535	0.00891	0.00831
	(0.05189)	(0.04768)	(0.05516)	(0.05171)
Green heart share	0.00103	0.00553	−0.01056	−0.00665
	(0.02246)	(0.02144)	(0.02302)	(0.02037)
Constant	0.54533**	0.58856**	0.56563**	0.70349**
	(0.27458)	(0.28355)	(0.26958)	(0.31864)
Control variables	Yes	Yes	Yes	Yes
Instrumented	Yes	Yes	Yes	Yes
Kleibergen-Paap statistic	18.34	41.05	17.68	36.05
Observations	672	672	512	512

Notes: (i) Included control variables are identical to the variables included in Tables 3.4 and 3.B.3. (ii) Unit of analysis is municipalities in 1990. (iii) Highway extent variables are in 1980 levels. (iv) Highway density is expressed in logarithm. (v) Possible endogeneity in the interaction terms is addressed following the single first-stage approach. (vi) Robust standard errors in parentheses. *$p<0.1$, **$p<0.05$, ***$p<0.01$

3.C First stage regression results

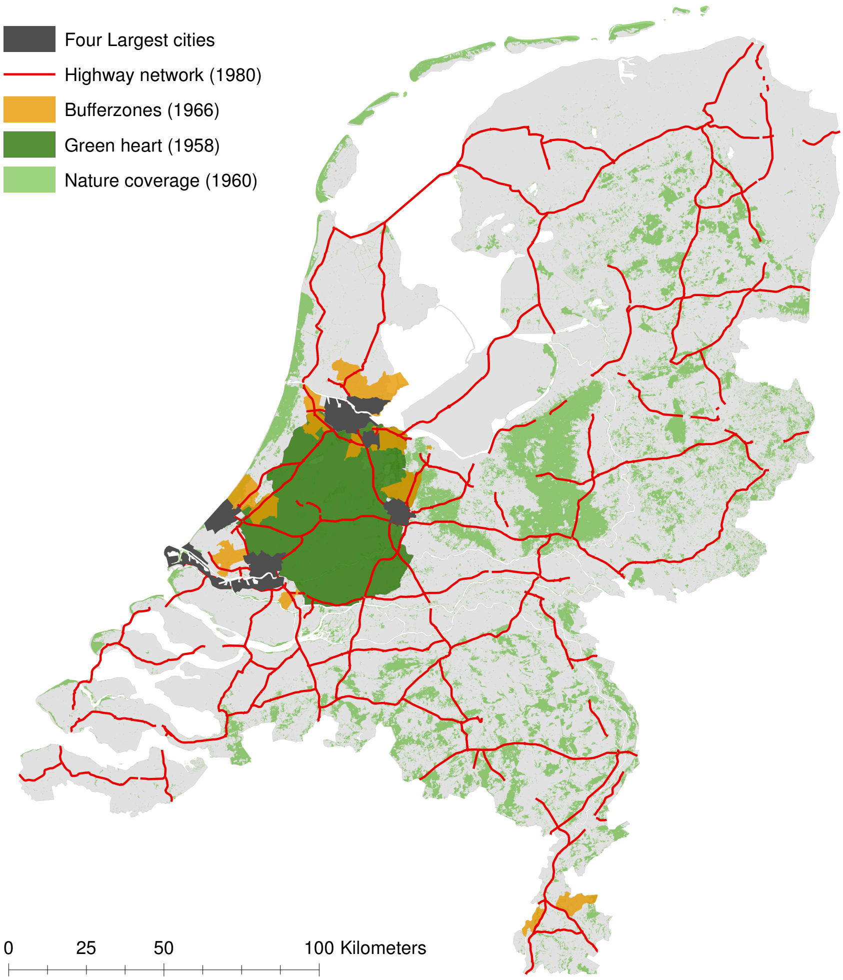
Table 3.C.1: Population growth and highways – Main results (first stage regressions)

	Main mode		Restricted specification	
Dependent variable:	Highway rays	Highway density	Highway rays	Highway density
	(1)	(2)	(3)	(4)
Road density	0.07009***	0.21781***	0.06987***	0.21516***
	(0.02117)	(0.03756)	(0.02109)	(0.03774)
Suburbs	0.09954	0.25874	0.09270	0.17375
	(0.18834)	(0.28181)	(0.18004)	(0.27025)
Central city	0.19029	1.38634*	0.09270	0.17375
	(0.52512)	(0.74724)	(0.18004)	(0.27025)
Buffer zone share	0.59057**	0.59485	0.58723**	0.55334
	(0.24998)	(0.43317)	(0.25015)	(0.42990)
Rail stations (1930)	−0.00130	0.06778	−0.00029	0.08026*
	(0.03073)	(0.04242)	(0.03087)	(0.04188)
Log. pop. (1930)	−0.21466**	−0.31256*	−0.21241***	−0.28463*
	(0.08316)	(0.16854)	(0.08197)	(0.15868)
Log. pop. (1960)	0.32202***	0.71421***	0.32359***	0.73374***
	(0.09229)	(0.16072)	(0.09053)	(0.15408)
Nature coverage share	−0.01797	−0.39332	−0.02032	−0.42251
	(0.34075)	(0.59755)	(0.33960)	(0.59908)
Green heart share	0.59021***	0.30454	0.58710***	0.26595
	(0.17269)	(0.23425)	(0.17428)	(0.23166)
Distance from Central city	−0.22875**	0.27024	−0.24055***	0.12359
	(0.11282)	(0.17408)	(0.08533)	(0.13320)
Reclaimed land share	−1.04179**	−0.83441	−1.02144**	−0.58159
	(0.50954)	(1.01588)	(0.49593)	(0.99425)
Constant	0.74610	−2.72208***	0.75110	−2.65994***
	(0.50657)	(0.88250)	(0.50785)	(0.87998)
R^2	0.096	0.127	0.097	0.125
Kleibergen-Paap statistic	10.96	33.62	10.97	32.50
Observations	811	811	811	811

Notes: (i) Highway extent variables are in 1970 levels. (ii) Highway density (1821, 1970) variables are expressed in logarithm, (iii) Robust standard errors in parentheses. $p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$

3.D Green heart, buffer zones and nature coverage

Figure 3.D.1: Green heart, buffer zones and nature coverage



Chapter 4

Spatial planning and segmentation of the land market: The case of the Netherlands

4.1 Introduction[1]

Many countries and cities engage in active land use policies, often by imposing restrictions. It has been argued in a recent literature that such restrictions can be harmful (Turner et al., 2014). When they are rigidly imposed, a serious distortion of the allocation of land can result, for instance by freezing an existing situation while economic forces call for substantial changes in land use. Although in many situations restrictions can be adjusted on request, and are therefore not as rigid as they appear to be, the associated costs are usually non-negligible. If the restrictions imposed on different types of land use differ, such zoning practices can lead to segmentation of the land market. For instance, if restrictions on residential land use are much tighter than those on commercial land use, the price of residential land can be significantly higher than that of commercial land on parcels that are close to each other and comparable in all other respects.

[1]This chapter is based on joint work with Jan Rouwendal and Lars Brugman. It is accepted for publication in *Land Economics*.
Financial support from ERC Advanced Grant OPTION (#246969) is gratefully acknowledged. I thank Ramona van Marwijk (Kadaster) for help with collecting and interpreting the data. I also thank Eric Koomen, and Netherlands Environmental Assessment Agency (PBL) for useful advices and spatial land use data.

A large recent literature has paid attention to the impact of land use restrictions on the prices of housing and residential land. However, there has been less attention for land use policy that relates to firms and their land use. Nevertheless, it is likely that there will be different implications of land use policy with respect to housing and firms. For instance, the homevoter or influential land user hypotheses (Fischel, 2001; Hilber and Robert-Nicoud, 2013) especially concerns residential land use. However, it is reasonable to expect that homeowners and landowners who want to increase the value of their property by restrictions on residential land use will also consider restrictions with respect to commercial land use.

In this chapter we analyze the effects of land development restrictions on the prices of undeveloped residential and commercial and agricultural lands in the Netherlands. The Netherlands is a densely populated country. The whole country is zoned and local land use plans determine which functions are allowed on a particular site. Often the land is designated exclusively either for agricultural, commercial or residential use. Urban growth is facilitated by the development of plans for (often) substantial extensions of existing urban areas. After acceptance by the local government, existing land use restrictions are lifted and the land is made ready for realization of the plans. However, while the responsibility for allocation decisions of commercial lands lies with local municipalities, the national government is actively involved in coordinating the development of new residential areas. As a result, there appear to be different regimes for residential and commercial development, where land use restrictions on the former are tighter than the latter.

If this planning process does not impose significant restrictions on urban development, one would expect that the price of undeveloped land, just outside existing urban areas, to be similar between different designated uses for parcels located close to each other. If the price of a certain land use would be higher, the familiar arbitrage mechanism would ensure that more areas would be converted to this use until provision is optimal and prices equate. However, if land use policy is restrictive, the prices of land of different uses can substantially differ. This observation is the focal point of the empirical analysis of this chapter.

We investigate how land development restrictions result in segmentation in the land market in the Netherlands. We start by examining median transaction values of undeveloped land designated for residential or commercial use in planned extensions of existing urban areas. If price differences are substantial, we conclude that land use policy does not only facilitate urban development, but also restricts it in comparison to the market outcome in the absence of such restrictions. We demonstrate that a significant difference exists between the median values of undeveloped residential and commercial lands, in the same local municipalities.

We proceed with an analysis of individual land transactions, taking on board a number of control variables reflecting possible sources of inter-municipal variation and using location fixed-effects to control for possible unobserved heterogeneity in local characteristics. The results of this more detailed investigation confirm that a substantial difference exists between the prices of undeveloped lands located close to each other, that it is significantly associated with land use designation, all else equal. We find that prices of undeveloped residential lands are approximately 88% higher than prices of industrial lands with similar characteristics and location. Next, we attempt to explain the difference by characteristics of these municipalities, that are potentially related to differences in their local land use policies such as accessibility to jobs, share of available agricultural land and the share of homeowners.

Finally, we extend our analysis to agricultural lands, as they form the main supply for urban development, and their values are therefore likely to be affected by land policies. Capozza and Helsley (1989) argue that the price of agricultural land in the vicinity of a growing city will reflect expectations with respect to future rents that will be realized after conversion to residential use. Land use restrictions will affect these expectations if they pose limitations on this conversion. In particular, agricultural land on which only agricultural activities are expected to be allowed for the foreseeable future will be of low value, even if it is located in the vicinity of a growing city. While Capozza and Helsley (1989) consider only residential land use at the edge of a growing city, the same logic implies that the prices of newly developed commercial and residential at locations close to each other to be equal, provided that differences in the conversion costs are properly taken into account. We conduct an additional analysis of agricultural land values, and find that agricultural land prices do not reflect expectations for future development, and that they are much lower than prices of nearby land that is ready to be developed either for residential or commercial purposes.

The chapter is organized as follows. In the next section we review relevant literature about the impact of land use regulation on land and house prices, and describe the planning system in the Netherlands. Section 4.3 discusses the data. Sections 4.4-4.7 contain our analysis of the land market. In section 4.4 we report the results of our analysis of land prices at the municipal level, using median transaction prices. In section 4.5 we report our analysis using information of individual transactions. Section 4.6 looks at explanatory variables. Section 4.7 contains the complementary analysis of agricultural land prices. Section 4.8 concludes.

4.2 Land use restrictions, land prices and the Dutch context

4.2.1 Land use planning and land prices

Recent literature has established that restrictions on land use can increase house prices by limiting the supply of residential land (Huang and Tang, 2012; Ihlanfeldt, 2007; Kok et al., 2014; Pollakowski and Wachter, 1990; Quigley and Raphael, 2005). An example is Glaeser et al. (2006), who find that zoning and heavy regulation makes housing supply relatively inelastic, which in turn restricts population growth and keeps housing prices and wages at high levels. Further evidence has shown that land regulation is strongly associated with lower supply of housing and reduction in residential construction (Green et al., 2005; Mayer and Somerville, 2000). While the effects of regulation on housing prices are generally found to be positive, the effects on land prices in urban areas can be negative. Ihlanfeldt (2007) studied the effects of land regulation on housing and land prices in cities in Florida and found that regulation increases housing prices but reduces land prices. This is explained by an increase in costs incurred by developers, in areas where development restrictions are more strict.

Hilber and Vermeulen (2016) show that a change in the planning system had a significant effect on housing construction and find a substantial impact of land use restrictions on house prices in the UK. Hamilton (1978) argues that municipal land regulation can generate monopolistic rent profits, as it exploits the relatively inelastic demand for labor (and consequently for housing) in an urban area.[2] Land use regulations can also be viewed as outcomes of political motives of local land owners. Fischel (2001) and Hilber and Robert-Nicoud (2013) argue that local owners of developed land benefit from stricter regulations on lands, and therefore they influence planning boards to increase these regulations.

A related possibility is that land use restrictions serve as a means to reduce urban sprawl. For instance, Saiz (2010) finds that land use regulation is stronger in areas that experience higher rates of demographic growth. Burchfield et al. (2006) argue that strict regulation imposed by municipalities deters developers and has a negative effect on residential development. They also show that – paradoxically - strict land regulations encourage developers to develop areas outside the municipal borders, where less regulation exists and development costs are lower. This suggests that the possibility to introduce local rigid land use restrictions may be limited by the size of the jurisdictions

[2] This idea is potentially relevant for the Netherlands where some municipalities have realized considerable benefits from being involved in land transactions.

introducing them. Indeed, Ihlanfeldt (2007) and Glaeser and Ward (2009) find that when land regulation is applied in urban areas with a large number of small local jurisdictions with similar levels of amenities, the effect of land regulation on house prices is small.

Turner et al. (2014) concentrate on differences in land prices occurring at jurisdictional boundaries to measure the impact of land use regulation. They find that a reduction of such regulation would result in significant welfare benefits, particularly for land owners in the edges of towns, where regulation is most restrictive for new development.

4.2.2 The Dutch context: preservation of agriculture, nature and open space

The Dutch spatial planning context differs from that in the US and the UK to which most existing literature refers. The Netherlands is a densely populated country, especially in the western core region (the *Randstad*). Restrictive land use regulations have been dominating the planning system since the end of WWII. The purpose of these regulations was to direct and accommodate the growing need for urban expansion while conserving the country's natural reserves and agricultural activity. The "Green Heart" (*Groene Hart*), a large agricultural area of approximately 2,400 square kilometer in the middle of the Randstad, was designated in the 1950's to ensure agricultural production in proximity to large population centers (Koomen et al., 2008).[3] Later on, the desire to preserve open (agricultural) space close to the country's largest population centers became the primary motive for keeping this part of the country dominated by green meadows. In 2004, the government defined an additional twenty rural and agricultural areas throughout the country as "National landscapes" (*Nationale Landschappen*, see map in appendix 4.A). The national landscape areas were defined in order to protect and preserve agricultural and cultural activity, as well as the environment and landscape in these areas, and therefore they impose heavy restrictions on residential and commercial development projects (Kuiper and De Regt, 2007). Additionally, following EU policy to preserve agricultural activity in naturally less-favored areas (LFA), such as peat meadows, river valleys and flood areas, approximately 225,000 hectares of agricultural lands in such areas are designated to remain agricultural in order to maintain the rural landscape and its biodiversity. Owners in such land were compensated by the government with approximately 94 Euro per hectare per year (Kuiper and De Regt, 2007).

Despite extensive land preservation policies, land use restrictions are not

[3] At that time the food shortages of the winter of 1945 were fresh in the planner's memories.

perfectly binding. To address the needs for residential and urban development, the planning system places an equally important attention to the facilitation of urban development. Land use regulations are updated every several years to accommodate changes in these needs. For example, in the course of time, the 'Green heart' shrunk to 1,800 square kilometers by 1993 (Koomen et al., 2008). Indeed, it has been argued that land use restrictions in the Netherlands are so flexible that in practice they hardly bind (Van Oort, 2004). This can be supported by the findings of Dekkers (2010), that prices of agricultural land in the vicinity of urban areas in the province of Noord-Holland reflect expectations that conversion and development will be allowed in the near future. However, these anticipation effects are also found to be strongly localized and therefore consistent with the view that in general existing restrictions are expected to keep in force.

One difficulty in the debate on the restrictiveness of Dutch land use restrictions is that it can be hard to demonstrate that actual land prices deviate from the counterfactual values in the absence of land use restriction. The study of market segmentation can contribute to this debate: if prices differ by (intended) land use, there must be binding restrictions on at least one type of land use.

4.2.3 The Dutch context: planning of urban extensions

Since the early 1990's, new residential development in the Netherlands is largely directed by the Fourth Memorandum for Spatial Planning Extra (abbreviated as 'VINEX'). Issued in 1991, the VINEX plan for new residential neighborhoods was conceived by the government in order to accommodate the housing needs of the growing population. It directed the construction of approximately 830,000 affordable housing units, of which 455,000 were planned in specifically designated "VINEX development areas" (Boeijenga et al., 2008; Broitman and Koomen, 2015; Koomen et al., 2008; Lörzing et al., 2006; Rietveld and Wagtendonk, 2004). The VINEX sites were located close to existing urban areas. To ensure the realization of these plans, the national governments had stimulated regional coordination through covenants by municipalities. Among other conditions, covenant agreement also depended on the requirement that employment centers would be constructed in short commuting distance from the new neighborhoods, in order to reduce car mobility (Boeijenga et al., 2008; Kruythoff and Teule, 1997).[4] The VINEX sites were thus agreed upon as the focal points of residential construction and outside these sites little residential development would take place.

[4]This difference is related to the fact that the Dutch constitution stipulates that housing (volkshuisvesting) is the responsibility of the national government.

The expected expansion of residential development, along with deregulation of the Dutch housing policy in the early 1990's,[5] increased the commercial interest in residential developments. The VINEX plan defined where the development would take place, restricted the development in other areas and directed most of the residential development to the owner-occupied sector. Therefore, investment in lands in the designated areas, still mainly agricultural at the time that the VINEX plan was published, offered low risk and high potential gains for commercial developers (Needham, 1997). Moreover the Dutch law stipulates that the owner of the land has the right to develop it[6] (within the limits imposed by the planning system) which implies that investments in land – at the right locations - by developers guaranteed them a strong market position when development would actually take place. Such investments occurred indeed on a large scale (Louw et al., 2003; Needham, 1997; Priemus and Louw, 2002). The price of residential land in VINEX areas was usually determined by the 'residual value method.' That is, the construction costs are subtracted from the expected market prices of the houses to be constructed. The developers buy the land from the municipality at this residual value, construct the houses and sell them. Construction only starts when a substantial part of the planned houses is sold to households,[7] which implies a strong downward rigidity of the prices of new houses (Vermeulen and Rouwendal, 2007).

In contrast with residential land, the responsibility for allocation of commercial land remained almost exclusively in the hands of the municipalities, which determine the allocation based on forecasts for regional employment demand (De Vor and De Groot, 2009; Louw and Bontekoning, 2007; Needham and Louw, 2006). There is a general impression of overprovision of commercial land by municipalities, which may be related to competition for employment. Between 1996–2010, areas designated for residential use grew from 213,700 ha to 231,400 ha (8.2% growth). During the same period, areas designated for commercial uses grew from 59,980 ha to 81,360 ha, a growth rate of almost 36% (Statistics Netherlands, 2015). Another notable regulatory difference with residential development is that the level of commercial land prices is monitored by the government and the chamber of commerce, who intervene in case of price increase (Needham and Louw, 2006).

[5]The deregulation of the housing market involved the withdrawal of government subsidies to newly developed social residential housing, and a shift of the responsibility for adequate housing from the central government to local municipalities and commercial developers (see more in Dieleman (1999); Priemus and Louw (2002); Vermeulen and Rouwendal (2007).

[6]Provided the owner is qualified to do so, as developers (of course) are, but farmers in general not.

[7]70 percent of the planned houses is often the rule of thumb.

4.2.4 The role of municipalities

While strategic spatial planning is determined by national and provincial level, municipalities are still a key player in the Dutch planning system and they have considerable market power within their own jurisdiction (De Vor and De Groot, 2009; Hajer and Zonneveld, 2000; Louw and Bontekoning, 2007; Louw et al., 2003; Needham, 1997; Van Oort, 2004). The most important element of spatial planning is the local land use plan ("*Bestemmingplan*") which is determined at the municipal level. Additionally, municipalities are the main suppliers of land for development (Louw and Bontekoning, 2007; Needham, 1997), and they are often actively involved in local land markets.

It is common practice that municipalities buy the land designated for urban extensions – which is usually in agricultural use[8] - from the original owners and sell it to developers after converting it to residential or commercial land.[9] Large scale extensions of urban areas are usually realized by a small number of commercial developers who cooperate closely with the municipalities. It is likely that such public-private partnerships are able to exercise some market power. In contrast, there is much less coordination and municipal involvement in the development of commercial areas. Firms develop their own land according to their specific needs. This, combined with the absence of similar national coordination, suggests that municipalities have less market power in commercial land development.

4.2.5 Hypotheses

The discussion in the previous subsections suggests that there are differences in the stringency of land use regulations with respect to residential and commercial development in the Netherlands. In particular, our discussion of Dutch planning practices suggests that residential land supply is more restricted than commercial land supply. This results in the conjecture that the price of undeveloped residential land will be higher than that of undeveloped commercial land. Put differently, the usual arbitrage mechanism that causes the price of land in nearby locations to be equal[10] is probably distorted. We expect that

[8]However, developers have also followed this strategy, which is all the more attractive to them since Dutch law stipulates that a land owner who is able to develop the land (within the limits of the local zoning plan) has the right to do so. Developers thus try to benefit from the increase in the value of land associated with the conversion to residential or industrial use and ensure themselves of new projects.

[9]The revenues are used to cover the servicing and development costs, but before the onset of the financial crisis in 2008 many municipalities realized large amounts of revenues from their land departments (Buitelaar, 2010; Louw et al., 2003; Needham, 1997).

[10]See, for instance, Albouy and Ehrlich (2013) for evidence that this mechanism works in U.S. cities.

the price of residential land to be the highest and that of agricultural land to be lowest. Failure to confirm this hypothesis would endorse the idea that land use restrictions are not actually binding.

4.3 Data

In the research and analysis we make use of unique data on residential and commercial land transactions, provided by Kadaster, the Dutch land registry.[11] The data includes 75,842 transactions of ready-to-be developed land between the years 2003 and 2013.[12] Residential land transactions form the majority of land transactions (71,578, or 94% of observations). 28,723 residential land transactions (or 40% of total residential transactions) are of plots in VINEX neighborhoods. The rest of the transactions are commercial, of which 3,683 are of land designated to industrial uses, 440 are of land for office use, and 141 are for retail use. Transactions per land use were identified based on deed registration and geographic location. Transaction deeds indicate the parcel's designated land use at the moment of sale.[13] In all transactions included, the buyer is a private household or firm.[14] We excluded transactions that did not only refer to land.[15] Designation for developed use in the cadastral data makes only the distinction between residential and commercial uses. To obtain a detailed information of commercial segments (industry, offices and retail) we compared the commercial transactions with land use information from

[11]There are not many papers that analyze land prices in urban areas. See Albouy and Ehrlich (2013) for a recent example. These authors do not find differences in land prices that are related to the type of land use (commercial vs residential).

[12]The decline in the number of residential land transactions after 2008 follows the global economic crisis, which affected new residential development in the Netherlands.

[13]Transactions of land converted to other uses soon after sale are still coded under their initial use at the time of sale. This suggests that if expectations exist for future conversion to other uses, they are expected to be reflected in the transaction price. This has a limited effect on the analysis of the values of non-agricultural land uses, mostly since conversion between residential and commercial lands uses is often prohibited. Furthermore, the analysis of agricultural land transactions utilizes this, as it is made under the direct assumption that their value depends on expectation for future conversion to other uses.

[14]Transactions in which the buyer is a public body or commercial developer were not included because conversion costs are perhaps not fully included in these transactions and large developers may have market power.

[15]Many transactions refer to land as well as a contract price (*aanneemsom*) for additional developments and the building to be constructed on the land. Identification of transactions that refer only to land was based on deed research and removal of groups of extreme positive price outliers, which often indicate that the transaction price includes additional irrelevant elements.

BAG (e, the Dutch register of addresses and buildings),[16] and with spatial information on business terrains from the Integral business areas information system (*Integraal Bedrijventerrein Informatie Systeem*, IBIS). For the agricultural land values analysis we make use of information about transactions of agricultural land, also provided by the land registry. In contrast to the data on newly developed residential and commercial sites, we have only aggregated information about agricultural land, regarding 26,629 agricultural land transactions in 60 agricultural regions (*Landbouwgebied*)[17] in the Netherlands between 2008–2013.[18]

Table 4.1 describes the median and mean annual land prices for lands of different uses.[19] It is evident that the values of lands designated for commercial uses are significantly lower. Between the years 2003–2013, transaction prices of industrial lands were on average 65% lower than residential land transactions. Despite a relatively low number of transactions, it is also apparent that prices per square meter of office and retail lands were also consistently lower than residential values. Moreover, extreme price differences exist between agricultural lands and other uses. While mean residential and industrial prices vary around €480 and €140 per square meter in 2013, the mean price of agricultural land in that period is approximately €4.95 per square meter (or €49,500 per hectare).

For the analysis of individual transactions we make use of transaction-specific location characteristics. Table 4.2 provides a summary of the chosen variables.

Spatial information for each transaction allows us to assign specific characteristics for each of the land transactions observed. Distance to nearest highway access point and to nearest train stations are included to control for variance in accessibility levels within municipalities, and were calculated us-

[16]Since the cadastral data includes only point coordinates of a land transaction (and not a delineated polygon), coupling of the transactions with the appropriate commercial segments was done by identifying the use of the nearest known structure, located up to a distance of 100 meters.

[17]The agricultural regions are an administrative division that covers the whole country, and often overlaps with municipal borders. Hence, one agricultural region can contain between 1 to 18 different municipalities, with an average of about nine municipalities for on agricultural region.

[18]In addition to limited availability for the years 2003–2007, the information regarding agricultural land transactions is also limited to region-totals of prices and hectare land sold, and it does not specify individual transaction data. Therefore, the figures presented here refer to mean land values for square meter, per year and agricultural regions.

[19]Median agricultural transaction prices cannot be computed since transaction data is only available for year and agricultural region aggregate. Our analysis focuses on median prices wherever possible, since they are less sensitive to transaction price outliers. However, for agricultural transactions we have only information about mean prices.

Table 4.1: Land transaction prices per land-use and year

Year	Residential			Industry			Offices			Retail			Agriculture		
	Median price/m2	Mean price/m2	N	Median price/m2	Mean price/m2	N	Median price/m2	Mean price/m2	N	Median price/m2	Mean price/m2	N	Median price/m2	Mean price/m2	N
2003	304.0	318.0	11,239	175.0	195.2	28	338.7	302.6	3	54.5	54.5	1	-	-	-
2004	333.6	358.8	13,099	148.9	210.9	132	248.3	255.1	25	348.4	293.9	7	-	-	-
2005	341.5	398.7	13,333	119.0	177.1	595	297.5	359.9	65	155.6	204.4	24	-	-	-
2006	377.2	430.6	11,839	122.2	211.2	640	233.1	286.0	82	104.9	168.2	25	-	-	-
2007	421.4	450.7	7,465	130.8	232.5	681	208.6	268.0	78	185.6	198.9	32	-	-	-
2008	497.1	527.2	4,286	118.6	218.2	622	297.5	311.4	87	170.0	242.5	20	-	4.36	5,383
2009	449.3	501.1	2,130	130.8	187.2	267	125.4	186.5	38	210.0	258.5	7	-	4.6	4,266
2010	535.8	561.8	2,989	124.9	157.2	266	182.5	195.8	18	96.0	173.6	7	-	4.52	3,854
2011	506.3	528.9	2,473	126.9	154.2	208	145.8	218.2	26	222.3	258.6	9	-	4.88	4,141
2012	501.8	524.5	1,445	135.7	168.9	143	194.1	204.2	8	521.7	440.8	3	-	4.71	4,241
2013	450.0	48.00	1,280	116.2	143.0	101	108.1	139.8	10	72.7	96.1	6	-	4.95	4,744
Total	371.5	415.5	71,578	125.5	198.3	3,683	210.1	276.0	440	155.5	210.6	141	-	4.67	26,629

Note: Data on agricultural land transaction is only available for the years 2008–2013.

Table 4.2: Summary statistics of individual land transactions characteristics

	Statistic	N	Mean	St. Dev.	Min	Max
All	Price/m2	75,842	403.75	222.09	25	1,300.00
	Distance to highway access point (meter)	75,842	3,126.27	2,796.27	95.51	28,221.57
	Distance to rail station (meter)	75,842	3,642.74	3,842.11	79.76	29,068.61
	Location within LFA	75,842	0.01	0.09	0	1
	Location within National landscape	75,842	0.08	0.28	0	1
Residential	Price/m2	71,578	415.49	218.39	75	1,300.00
	Distance to highway access point (meter)	71,578	3,070.89	2,682.98	143.14	28,221.57
	Distance to rail station (meter)	71,578	3,515.14	3,700.45	79.76	28,388.44
	Location within LFA	71,578	0.01	0.09	0	1
	Location within National landscape	71,578	0.08	0.27	0	1
Industrial	Price/m2	3,683	198.31	185.23	25	885.55
	Distance to highway access point (meter)	3,683	4,199.64	4,286.41	95.51	27,815.48
	Distance to rail station (meter)	3,683	5,973.25	5,409.01	101.02	29,068.61
	Location within LFA	3,683	0.01	0.08	0	1
	Location within National landscape	3,683	0.14	0.35	0	1
Offices	Price/m2	440	276.01	204.85	25.91	875.66
	Distance to highway access point (meter)	440	3,006.94	2,835.46	175.54	16,159.35
	Distance to rail station (meter)	440	4,644.38	3,940.23	186.15	22,696.75
	Location within LFA	440	0.02	0.14	0	1
	Location within National landscape	440	0.14	0.34	0	1
Retail	Price/m2	141	210.55	174.77	26	799.00
	Distance to highway access point (meter)	141	3,574.09	3,865.68	99.38	17,843.73
	Distance to rail station (meter)	141	4,418.25	4,766.99	132.97	20,273.88
	Location within LFA	141	0.01	0.08	0	1
	Location within National landscape	141	0.12	0.33	0	1

ing GIS, based on information from NWB (*Nationaal Wegen Bestand*), the Dutch road and rail network data (*Rijkswaterstaat*, The Dutch ministry of infrastructure and environment, 2013, 2014). National landscapes are defined in the Nota Ruimte planning memorandum (Dutch Ministry of infrastructure and environment (VROM, 2004)). and were made available by Koomen et al. (2008), who also provided the spatial information for the VINEX neighborhoods (VROM, 1993) and the LFA areas (Kuiper and De Regt, 2007; MNP, 2005; LNV, 2008).

For the municipality level analysis we make use of information on municipal amenities and political characteristics. Table 4.3 provides a summary of the variables chosen.

To measure accessibility at the municipal level, we use the number of jobs (in logarithm) which are accessible within a commuting time of 60 minutes by road from each municipality's population center. Travel time is calculated based on NWB (*Rijkswaterstaat*, 2015).[20]

Information on the share of agricultural land in each municipality in 2010, as well as the share of homeowner households in each municipality, is available from statistics Netherlands. Voters' data for local municipalities' council in 2010 was available from *Databank Verkiezingsuitslagen*.[21]

Table 4.3: Summary statistics of municipal characteristics variables

Statistic	N	Mean	St. Dev.	Min	Max
Accessibility to jobs	1,005	7.040	0.999	0.191	8.552
Share of agricultural land (2010)	1,005	0.535	0.237	0	0.940
Share of homeowner households	889	0.649	0.078	0.408	0.817
Share of voters to GroenLinks (2010)	1,005	0.056	0.056	0	0.320
Share of voters to VVD (2010)	1,005	0.144	0.072	0	0.411
Area included in National landscape (share)	1,005	0.203	0.328	0	1
Share of municipal area include in the GH (1993)	1,005	0.042	0.183	0	1
VINEX municipality dummy	1,005	0.221	0.415	0	1

[20]Travel time in different roads was assigned according to the common maximum travel speed in highways and main roads (*Rijkswaterstaat*, 2015) - Travel speed of 100 km/h was assigned in highways ("A-roads"), and 50 km/h for main provincial roads ("N-Roads"). Where information was unavailable, a traffic speed of 50 km/h was assumed for non-urban provincial roads (Gutiérrez et al., 2010) and 30 km/h for urban roads (Rietveld and Bruinsma, 1998). Inner-municipal travel time (internal travel time) is computed following Koopmans et al. (2012), based on municipality area and internal travel speed of 30 km/h.

[21]Source: http://www.verkiezingsuitslagen.nl/ [accessed : November 28, 2016]

4.4 Analysis of median prices

To study the presence of land market segmentation we begin by testing the equality of prices of undeveloped commercial and residential lands, at the municipal level, by estimating the following model:

$$P_{j,t}^{R} = \alpha + \beta P_{j,t}^{I} + \sum_{t} \beta_t Y_t + \epsilon_{j,t}, \tag{4.1}$$

where $P_{j,t}^{R}$ is the median price per square meter of newly developed residential land in municipality j and time t and $P_{j,t}^{I}$ the median price per square meter of industrial land in j and time t. Land transactions of parcels designated for retail and office uses are excluded from this analysis due to a relatively low number of observations.[22] If land rents are equal for both residential and industrial uses, we should observe that $\alpha = 0$ and $\beta = 1$ in the absence of conversion costs. If conversions costs are present α should reflect the difference in conversion costs between residential and industrial land, in a free market equilibrium.

The estimation results presented in Table 4.4 show α and β coefficients significantly different from the respective null hypothesis values of 0 and 1.[23] This indicates that we can reject the hypothesis that values of newly developed land, and hence also its rents, are equal for both residential and industrial uses. The model's constant term α is positive and its value is approximately €220. According to Buitelaar and Witte (2011), the conversion costs of industrial land are approximately €21 per square meter lower than the conversion costs of residential or mixed land use. The large difference between the land values thus cannot be attributed to differences in conversion costs alone. Although the value of β is estimated to be smaller than 1, the relatively high value of α ensures that the transaction price per square meter of residential lands is substantially higher than the price of industrial lands. We thus find that the prices of newly developed residential and industrial land are often substantially different at the municipal level.

4.5 Analysis of individual land transactions

In this section we deepen our analysis by looking at individual land transactions to identify the difference in values of undeveloped lands of different uses.

[22] Transactions of all office, retail, industry and residential land uses are only observed in 14 municipalities and year combinations, compared with 1,005 municipality-year combinations in which only industrial and residential transactions are observed.

[23] F-test results for the null hypotheses $\alpha = 0$ and $\beta = 1$ both show p-values lower than $Pr(F) < 0.0000$.

Table 4.4: Estimation results for the equality of the prices of residential and industrial land

	Residential price/m2 (1)
Industrial price/m2	0.611***
	(0.0413)
Constant	220.4***
	(47.17)
Observations	1,005
Adjusted R^2	0.257

Notes: (i) Year dummies are included. (ii) Standard errors are clustered for each municipality. (iii) Standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01

This has several advantages. One is that it allows us to include characteristics of individual parcels, and to control for intra-municipal differences in accessibility, or other local differences in the productivity of land. The heterogeneity between transactions within a given municipality may be related to the fact that planners have knowledge about basic factors influencing the productivity of land, and use this when allocating land to specific uses. Information on individual transactions offers better control for such variation than median prices do.

A second reason is that it is well-known that parcel size is negatively related to the price per square meter of land. Since land designated for commercial use is usually traded in larger parcels than land designated for residential use, this may well bias results based on median prices per square meter. Third, the use of individual transaction data allows us to distinguish between values of lands of different commercial segments, and to utilize the relatively small number of transactions referring to offices and retail.

We estimate the following model based on the full sample of residential, industrial, offices and retail transactions:

$$ln(P_{i,t}) = \beta_0 + \sum_k \beta_k L_i^k + \sum_l \beta_l Z_{i,l} + \sum_t \beta_t Y_t + \sum_j \beta_j R_j + \epsilon_{i,t} \qquad (4.2)$$

where $ln(P_{i,t})$is the transaction price per square meter in logarithm of land parcel i, L_i^k is a vector of $k = [1, 2, 3]$ dummy variables which is equal to one if the transaction is of land use k (residential, office or retail), and equals to zero otherwise. Industrial transactions form the base group. the corresponding coefficient β_k is interpreted as the inherent difference in value of undeveloped residential land, compared with the value of undeveloped industrial land in the same municipality and year. $Z_{i,l}$ are control variables, including an indicator for the soil type of the parcel sold in transaction i (we control for 14 soil types), indicator variable for whether the parcel is located within a national landscape area and an indicator for location within a less-favored agricultural area. Y_t and R_j are dummies indicating year t and municipality j, of land transaction i.

Although the approach outlined above allows us to control for many effects that potentially bias our results, one may still be concerned about unobserved variables that may have the same effect. To address this issue, we carry out an additional fixed-effects analysis. That is, we compare only transactions of parcels that are located close to each other by introducing dummy variables. The matched transactions refer to land parcels located within 500m of each other. This resulted in some loss of observations, but in our data undeveloped parcels of land of different uses are often adjacent to each other.[24] The threshold of 500m was chosen based on testing a number of relatively small distances. To demonstrate the robustness of the results we also report estimation results using several other distance threshold values (see Appendix 4.B). We estimate both models with municipality clustered standard errors to account for possible correlation in the model's residuals.

Estimation results are presented in Table 4.5. Columns 1 and 2 refer to OLS, columns 3 and 4 to the spatial matching fixed-effects analysis. Column 1 shows a coefficient of 0.713 for residential lands, which implies that the price per square meter of undeveloped land designated for residential use is approximately 104% higher than industrial lands sold in the same municipality and year. The matching analysis returns a somewhat smaller coefficient of the same order of magnitude, which suggests a price gap of 88%. We regard this as the main result that establishes the segmentation of the land market due to restrictive land use policy.

In columns 2 and 4 we allow for differences in the premium for residential land between VINEX and other locations. As we expected, the premium is lower in such locations, but the differences are relatively small and statistically insignificant. Also in the VINEX locations there is a substantial price gap between land designated for residential and industrial use.

Table 4.5 also reveals that there are differences between the types of commercial land use. Land designated for offices sells at a 13-20% higher price per

[24]Matched parcels can be located in different municipalities.

Table 4.5: Estimation results: The divergence in prices of residential and industrial land

	OLS (Full sample)		Spatial matching FE: Distance threshold (500m)	
	Log. price/m2	Log. price/m2	Log. price/m2	Log. price/m2
	(1)	(2)	(3)	(4)
Residential	0.713***		0.633***	
	(0.0310)		(0.0515)	
Residential (not VINEX)		0.720***		0.644***
		(0.0308)		(0.0492)
Residential (VINEX)		0.679***		0.587***
		(0.0355)		(0.0832)
Offices	0.126***	0.123***	0.200**	0.195**
	(0.0430)	(0.0424)	(0.0802)	(0.0811)
Retail	0.130*	0.129*	−0.0601	−0.0574
	(0.0701)	(0.0700)	(0.0923)	(0.0910)
Log. Parcel size (Sq. meter)	−0.144***	−0.144***	−0.141***	−0.141***
	(0.0104)	(0.0104)	(0.0232)	(0.0231)
Log. distance to highway ramp	0.0189	0.0185	−0.0180	−0.0138
	(0.0151)	(0.0149)	(0.0641)	(0.0658)
Log. distance to railway station	−0.0310*	−0.0315*	−0.0863	−0.0871
	(0.0158)	(0.0160)	(0.0666)	(0.0678)
Location within LFA	0.0658	0.0690		
	(0.0534)	(0.0529)		
Location within nat. landscape	0.00574	0.00388	−0.112	−0.115
	(0.0300)	(0.0302)	(0.154)	(0.153)
Constant	5.084***	5.082***	6.592***	6.572***
	(0.177)	(0.175)	(0.913)	(0.922)
Observations	75,842	75,842	7,969	7,969
Adjusted R^2	0.729	0.730	0.320	0.320
Fixed-effects	No	No	Yes	Yes
Fixed-effects Groups			290	290

Notes: (i) Year, municipality and soil type dummies are included. (ii) Standard errors of the OLS estimations are clustered for each municipality. (iii) Standard errors in parentheses. $^{*}p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$

square meter than land designated for industrial use. The coefficient for the retail dummy is insignificant. Note, however, that there are only a handful of such transactions in our data.

As we expected, parcel size has a significant negative effect on the price per square meter. Somewhat surprisingly, the coefficients on the accessibility variables – distance to highway ramp and railway station – are insignificant. The same is true for the coefficients for the dummies indicating the location of a parcel in a less favored areas (LFA's) or national landscape.

The differences between the OLS and matching analysis are limited, but non-negligible. Additional sensitivity analysis in which the threshold is set to 250, 400, 600 and 1000 meters shows that the estimated residential dummies coefficients are robust and maintain similar signs, values and significance level (see appendix 4.B).

4.6 Explaining the divergence in land values using municipal characteristics

The divergence in land values is likely to vary between regions and may be related to local circumstances. To address this we investigate in the present section several possible explanations for the local price gap, using the difference between the median price of residential and industrial land as a summary measure of this gap.[25] More specifically, we estimate a model in which we explain the difference between median prices per square meter of residential and industrial lands by differences in local amenities and political characteristics:

$$ln(P^R_{j,t}) - ln(P^I_{j,t}) = \beta_0 + \sum_m \delta_m X_{m,j} + \sum_t \beta_t Y_t + \beta_u u_j + \epsilon_{j,t} \qquad (4.3)$$

where $ln(P^R_{j,t}) - ln(P^I_{j,t})$ is the difference between the median prices of residential and industrial lands in municipality j and year t (in logarithm), $X_{m,j}$ are control variables, which will be further discussed below. Y_t are year dummies, and u_j indicates the share of municipal area which is undeveloped in 2010, whether in agricultural use, or as an open space (undeveloped area excluding agricultural activities).

Estimation results are presented in Table 4.6. Column 1 refers to the regression of the price difference, whereas columns 2 and 3 give the results referring to both components of the difference. The R^2 of the equation referring

[25] Annual median transaction prices of office and retail lands are again excluded due to low number of observations.

to the price gap is much lower than that of the two other equations, due to the fact that the error term is a much larger component of the variation in the difference. The reason is that the prices of both types of land use are partly explained by the same factors, which lose their impact if attention shifts to the difference between the two prices.

We use the control variables $X_{m,j}$ to investigate a number of possible explanations of the price gap between residential and commercial land. The first is that accessibility to jobs is important. The reason could be that municipalities located in the economic center of the country, where population and job density are high, can more easily afford to charge higher prices for industrial as well as residential land than those located in more peripheral areas. The significantly negative coefficient in column 1 shows that the data offers support to this hypothesis. One percent increase in the number of jobs accessible within one hour of commuting from the center of a municipality reduces the land value gap by approximately 0.11 percent. Note also that the impact of accessibility on the price of industrial land is stronger than that on residential land.

A second hypothesis is that municipalities with a large amount of agricultural land show a larger gap between the prices of residential and commercial land since the supply of the former is restricted by national plans, whereas that of the latter can be determined by the municipality. In agricultural areas local employment is often a source of concern and municipalities may therefore be inclined to keep the price of industrial land low so as to attract non-agricultural firms. This second hypothesis is also confirmed by the data. We find a large positive impact of the share of agricultural land in the municipality, which is caused by the low prices of commercial land.

Following Fischel (2001) and Hilber and Robert-Nicoud (2013), one may conjecture that development restrictions may be the result of a political motive and are influenced by local voters. If this holds, we expect that municipalities with a higher share of homeowners would result in tighter land development regulations, which would reflect in higher land values. These restrictions are likely to affect values of both land uses, which may increase or decrease the gap in values. However, national coordination in residential development may limit the possibilities for homevoters or influential landowners to affect local land prices. To test the homevoter hypothesis, we included the share of homeowners per municipality in the regression. We find a significantly negative coefficient for the price gap and an insignificant positive coefficient on the price of residential land. This suggests a rejection of the homevoter hypothesis. However, the higher price for industrial land in municipalities with many homeowners may indicate a 'NIMBY' effect in which owner-occupiers

Table 4.6: Estimation results: Explaining the gap in the values of different land uses

	Log. gap Residential - Industrial price/m2	Log. Residential price/m2	Log. Industrial price/m2
	(1)	(2)	(3)
Log. Accessibility to jobs	−0.111***	0.243***	0.354***
	(0.0265)	(0.0402)	(0.0532)
Share of agricultural land (2010)	0.598***	−0.581***	−1.179***
	(0.129)	(0.133)	(0.201)
Share of homeowner households	−1.184***	0.302	1.485***
	(0.380)	(0.297)	(0.468)
Share of voters to GroenLinks	0.147	0.390	0.242
	(0.423)	(0.352)	(0.430)
Share of voters to VVD	0.453	0.569*	0.116
	(0.353)	(0.291)	(0.370)
Area included in nat. landscape (share)	−0.0102	0.173**	0.184**
	(0.0720)	(0.0680)	(0.0858)
Share of municipal area include in the GH (1993)	−0.0353	0.320***	0.355**
	(0.146)	(0.112)	(0.154)
VINEX municipality	0.0255	0.168***	0.143*
	(0.0613)	(0.0513)	(0.0737)
Constant	1.979***	4.020***	2.041***
	(0.313)	(0.320)	(0.467)
Observations	889	889	889
Adjusted R^2	0.073	0.481	0.428

Notes: (i) Year dummies are included. (ii) Standard errors are clustered for each municipality. (iii) Standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01

try to protect their housing wealth against possible negative external effects of industrial land use.

The next two variables refer to the share of votes in the municipal council elections of 2010 of two political parties: GroenLinks, the Dutch green political party, and VVD, the liberal party.[26] These parties are positioned in opposite sides of the political map. The green party supports development regulations and protection of nature. The liberal party supports entrepreneurs and reduced regulation. GroenLinks is a relatively small party (average election result in Dutch municipalities in 2010 is 5.6%), while VVD is a leading political party (average result is 14.4%).[27] The correlation between the shares of the two parties is relatively weak ($\rho = 0.23$).[28] The estimation results show no evidence that the price gap is influenced by local support for political parties' activities. We find a weakly significant effect on the price of residential land, which is perhaps related to homevoter behavior as owner-occupiers are overrepresented among VVD voters.

The last three variables included in the regression refer to national land use planning. The first two of them, the share of the municipality included in a National Landscape area and the share included in the green heart (GH) of the Randstad, restrict all development in areas that are judged to be of special value. As expected, they increase the prices of residential as well as industrial land, while there is no specific impact on the gap between the two prices.

Finally, we included a dummy that indicates if a municipality has a VINEX agreement (namely, includes a VINEX location within its boundaries). Again, there is no impact on the price gap. It is, however, surprising that a VINEX agreement has a positive effect on both residential and industrial, since the larger supply of both types of land in such municipalities is expected to lead to the opposite result. A possible explanation is that house types in VINEX areas tend to be more valuable compared to other residential development areas. Moreover, it may be that higher industrial land values reflect expectations for higher future accessibility of workers.

Summarizing our findings, the price gap that is the focus of this chapter

[26] Dutch municipal council election results are available online from the Dutch Electoral Council website at www.kiesraad.nl [Accessed: April 25th 2016]. In 41 out of 431 municipalities, GroenLinks participated in the 2010 municipal council elections as part of a coalition with other left wing parties. GroenLinks ideology is available at: https://groenlinks.nl/ [Accessed: December 3rd 2016]. VVD ideology is available at https://www.vvd.nl/standpunten [Accessed: December 3rd 2016].

[27] Note that the Dutch political landscape is extremely fragmented in comparison to that in many other countries.

[28] Moreover, the correlation between the share of homeowners and the share of voters to the Green political party and votes to VVD is moderate and weak, respectively ($\rho_{homeowner,GL} = -0.30$, $\rho_{homeowner,VVD} = -0.065$).

is associated with job accessibility and the presence of agricultural land. The negative relationship with the share of homeowners is unexpected, and may be related to NIMBY-behavior that could be generated by similar sentiments as homevoter or influential landowner behavior.

4.7 Analysis of agricultural prices

Agricultural land values are a key component in the values of other land uses, as they reflect the pre-development value of the land, and its alternative use in most cases (Capozza and Helsley, 1989, 1990). If the expectation for future development exists, one would not expect to find a difference in values of undeveloped lands designated to different use, but if land use restrictions are expected to persist in the future, the value of the nearby agricultural land will be considerably lower. In the foregoing section we have provided evidence that land use restrictions in the Netherlands are tight enough to cause segmentation of the markets for newly developed residential and industrial land. In the Netherlands conversion of agricultural land to residential or industrial use is as a rule prohibited. The analysis of the previous sections suggests that such restrictions may also be binding and at least partly prevent the agricultural land close to cities to reflect its future value, after conversion, through the mechanism described by Capozza and Helsley (1989). To investigate this, we estimate the following model:

$$ln(P^A_{j,t}) = \beta_0 + \beta_R ln(P^R_{j,t}) + \beta_I ln(P^I_{j,t}) + \sum_l \beta_l Z_{j,l} + \sum_t \beta_t Y_t + \epsilon_{j,t} \quad (4.4)$$

where $ln(P^A_{j,t})$ is the mean price (in log) of agricultural land in year t in agricultural region j, $ln(P^R_{j,t})$ is the mean price (in log) of residential land in year t and region j, $ln(P^I_{j,t})$ is the mean price (in log) of industrial land in year t and region j. Since information on agricultural land transactions is available only as year and agricultural region aggregate, all independent variables (including residential and industrial prices) also reflect mean values per year and region in order to maintain consistency. As in the analysis in 4.1 and 4.2, we exclude office and retail transactions due to a relatively low number of observations. Y_t are year dummies and $Z_{j,l}$ are control variables of land characteristics in each agricultural region, among which are the share of area included in national landscape area, share of areas defined as LFA area, nature coverage in 2008, and prevalence of each of the 14 most common soil types in the Netherlands. We estimate the model with agricultural region clustered standard errors to account for possible correlation in the model's residuals.[29]

Table 4.7 presents the results of model 4.4. Positive and significant coefficients of residential and industrial land values suggest that in municipalities where residential and industrial values are higher by 1%, the price of agricultural land tends to be higher by 0.15% and 0.051% respectively. This corresponds with the findings of Dekkers (2010).

Table 4.7: Analysis of agricultural land values

	Log. agricultural price/m2 (1)
Log. residential price/m2	0.150***
	(0.0521)
Log. industrial price/m2	0.0514**
	(0.0257)
Area included in nat. landscape (share)	−0.000258
	(0.00136)
Less-favored agricultural area (share)	−0.00515*
	(0.00306)
Nature coverage 2008 (share)	−0.00325
	(0.00277)
Constant	0.546
	(0.403)
Observations	281
Adjusted R^2	0.537

Notes: (i) All variables (including agricultural, residential and industrial land transactions) reflect mean values per year and agricultural region. (ii) Year dummies and soil type prevalence controls are included (iii) Standard errors are clustered for each agricultural region (iv) Standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01

These findings suggest that agricultural land values are higher in areas that experience higher industrial and residential price levels, possibly due to expectations for future land development. These expectations can be either for a

[29] Suspicion of reversed causality between the prices of agricultural lands and of residential and industrial uses are reduced when supply restrictions in the land market are considered. It is therefore less probable that variance in the value of agricultural lands has a substantial effect on land values of other uses.

municipal decision to convert lands to industrial uses, or for a possible eventual policy change and relaxation of residential development restrictions. However, although the estimated effects are statistically significant, the magnitude of the effect on the price of agricultural land remains negligible compared with the values of undeveloped land designated for future development. Therefore, even if evidence of expectations for land conversion exists, it is hardly capitalized in the values of agricultural lands. The probable explanation is that expectations with respect to future conversions are limited to specific areas close to existing cities, while our data refer to medians of relatively large areas. Nevertheless, it is noteworthy that even in the Randstad area, where cities and villages are located close to each other, there remains a huge difference between the average price of agricultural land and that of ready-to-be built upon residential and commercial land. Without restrictive land use planning, expectations for future urban development would probably drive up the price of all land uses.

4.8 Conclusion

Our analysis of the transaction prices of lands reveals a clear pattern: undeveloped residential land has the highest price and agricultural land the lowest. The price difference between undeveloped residential and commercial lands, industrial land in particular, as well as the price difference between agricultural and undeveloped industrial lands, cannot be ascribed to differences in conversion costs. We find that undeveloped lands designated for residential use are valued more than twice as high as land designated to industrial lands, all else equal. Moreover, the gap does not disappear if we introduce a number of controls or restrict the comparison between parcels located within a few hundred meters of each other. This shows that the planning restrictions are binding and prevent the functioning of the usual arbitrage mechanism. We show that this gap in values is influenced by local land policies, homeownership rates and accessibility levels.

We thus find that land use policy in the Netherlands is restrictive in at least two ways: residential development is more limited than commercial development and land that is zoned to be exclusively agricultural is in general expected to be agricultural for the foreseeable future.

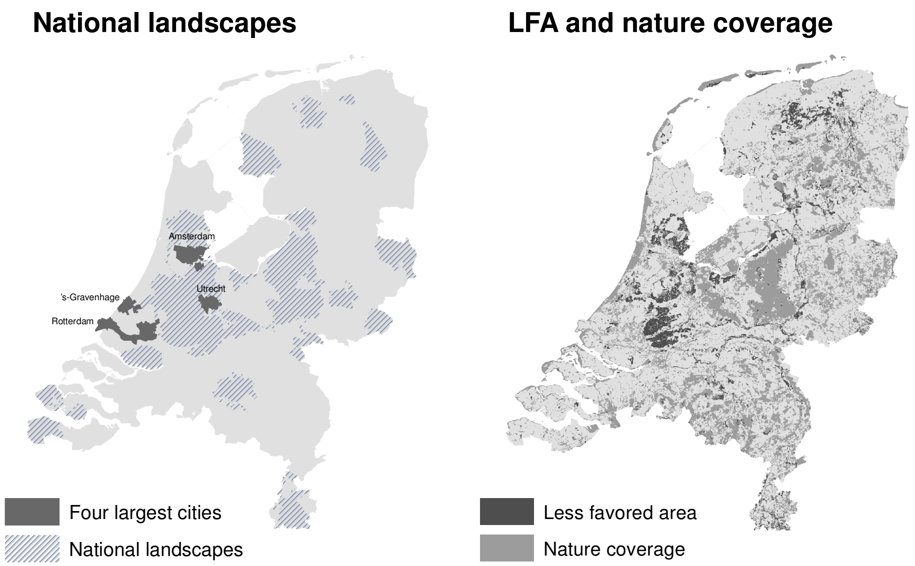
Our analysis does not directly reveal any costs or benefits of spatial planning. Benefits, for instance in the form of protecting open space, surely exist, although it may be doubted if the high price of urban residential land can be justified by the protection of open space that results from it (Vermeulen and Rouwendal, 2007). Costs are also present. For instance, it is hard to see why open space should be protected by tightening the supply of residential land much more than that of industrial land. The large difference between the prices for undeveloped residential and industrial land signals misallocation. If one thinks that the supply of commercial land is appropriate, then not enough land for residential use is provided. Conversely, if the level of residential land prices is appropriate for maximizing the benefits of land use planning, the supply of commercial land should consequently be tightened so as to bring its price at a comparable value.[30]

If the price of undeveloped residential land is higher than optimal from a social welfare maximizing point of view, households pay too much for their houses. Since the price of undeveloped land is reflected in the price of all housing, additional supply of residential land could have resulted in house prices following the development of construction cost, which would have resulted in better affordability and a substantial benefit for many home buyers. A lower price of undeveloped residential land might have also been reflected in the housing type composition of new neighborhoods, and construction in lower density should be expected (Ahlfeldt and McMillen, 2014; Epple et al., 2010).

[30] As far as we know there is no support for this point of view in the Netherlands, which may be interpreted as suggesting that the price of residential land is too high.

4.A Land development restrictions - National landscapes and less-favored areas (LFA)

Figure 4.A.1: Land development restrictions - National landscapes and Less-Favored Areas (LFA)



4.B Estimation results: Testing the divergence in prices of residential and industrial land (robustness check with different threshold subsamples)

Table 4.B.1: Estimation results: The divergence in prices of residential and industrial land (different threshold subsamples)

	Spatial matching FE: Distance threshold (250m) Log. price/m2	Spatial matching FE: Distance threshold (400m) Log. price/m2	Spatial matching FE: Distance threshold (500m) Log. price/m2	Spatial matching FE: Distance threshold (600m) Log. price/m2	Spatial matching FE: Distance threshold (1000m) Log. price/m2
	(1)	(2)	(3)	(4)	(5)
Residential	0.554***	0.620***	0.633***	0.578***	0.658***
	(0.0990)	(0.0535)	(0.0515)	(0.0553)	(0.0438)
Offices	0.251*	0.235***	0.200**	0.132	0.101
	(0.129)	(0.0894)	(0.0802)	(0.0833)	(0.0621)
Retail	−0.00811	−0.0299	−0.0601	−0.0348	0.0148
	(0.154)	(0.108)	(0.0923)	(0.101)	(0.0836)
Log. Parcel size (Sq. meter)	−0.185***	−0.148***	−0.141***	−0.159***	−0.155***
	(0.0536)	(0.0285)	(0.0232)	(0.0196)	(0.0164)
Log. distance to highway ramp	0.111	−0.00512	−0.0180	−0.0277	0.00509
	(0.220)	(0.0925)	(0.0641)	(0.0621)	(0.0374)
Log. distance to railway station	−0.171	−0.189**	−0.0863	−0.0510	−0.0287
	(0.156)	(0.0901)	(0.0666)	(0.0601)	(0.0402)
Location within LFA				0.124	−0.0722*
				(0.159)	(0.0398)
Location within nat. landscape			−0.112	−0.000289	0.168**
			(0.154)	(0.195)	(0.0841)
Constant	6.749***	7.354***	6.592***	6.622***	6.024***
	(2.038)	(1.147)	(0.913)	(0.877)	(0.531)
Observations	1,814	5,251	7,969	11,284	26,813
Adjusted R^2	0.447	0.332	0.320	0.314	0.328
Fixed-effects Groups	112	227	290	343	490

Notes: (i) Year, municipality and soil type dummies are included. (ii) Standard errors are clustered for each municipality. (iii) Standard errors in parentheses.
*p<0.1, **p<0.05, ***p<0.01

Chapter 5

Estimating the value of proximity to water, when ceteris really is paribus

5.1 Introduction[1]

The purpose of hedonic analysis is often to measure the marginal willingness to pay for a public good or an external effect such as the proximity to a park or aircraft noise. This is usually complicated by the extreme heterogeneity of the housing stock. In general, it is difficult to find identical houses and even if there are no differences in observed characteristics, there can be considerable differences in unobserved attributes. These unobserved attributes are a potential cause of omitted variable bias in the coefficients of the hedonic price function and in measures of the marginal willingness to pay based on them.

In this chapter we explore a way to circumvent this problem by concentrating on newly developed residential areas in the Netherlands. It is often the case that a limited number of dwelling types are made available on these sites and that a (possibly large) number of each type is constructed. The technical attributes of units of the same type are identical, and this means that a major source of unobserved heterogeneity is absent. Moreover, these houses are lo-

[1]This chapter is based on joint work with Jan Rouwendal (VU University Amsterdam and Tinbergen Institute, The Netherlands) and Ramona van Marwijk (Kadaster, the Dutch land registry). Apart from minor changes, this chapter is published as Rouwendal, Jan, Levkovich, Or and Ramona Van Marwijk. *"Estimating the Value of Proximity to Water, When Ceteris Really Is Paribus."* Real Estate Economics (2016).

I would like to thank Jos van Ommeren, Hans Koster, Fred Huibers, anonymous referees and the editor, Ed Coulson for their useful and insightful comments. I also thank Ed Coulson for suggesting the current title of this chpater.

cated close to each other, implying that neighborhood characteristics are also equal. To reduce the risk that houses that were initially identical have become different because of later investments, we restrict the analysis to houses that were constructed less than 15 years before the transaction prices were observed. These houses, which are identical in so many respects, may nevertheless differ in their proximity to water and this gives us the opportunity to measure the impact of this amenity under conditions that are close to ideal.

Reducing the estimator bias and revealing the price and willingness-to-pay patterns for proximity to water can be useful for planning and policy-making about future residential areas. Since lakes and waterways are often developed in order to increase the attractiveness of a new neighborhood, overestimation of the value of canals and lakes in residential areas may have encouraged planners to favor them over the development of other neighborhood amenities, which may have been more valuable in the eyes of residents and homeowners (including other environmental amenities). Moreover, this research aims to further demonstrate how unobserved heterogeneity can cause a bias in the estimated effect of non-market housing attributes, as reflected in the results of hedonic models.

5.2 Effects measured in the previous literature, and the methods that were used

Several researchers have applied hedonic models to measure the effect of water on residential property values and most of them have found value-increasing effects of water on house prices.

Lansford and Jones (1995) used hedonic regression to measure the effects of proximity to water in the Colorado River basin (Texas). For the estimation of the hedonic regression, they used a Box–Cox transformation of the dependent variable and of the continuous distance variable. They also assumed that proximity to water may have very little influence on properties which are located over 4,000 feet from the water, and therefor included a dummy variable to capture the effect of proximity to water in such distance ranges. They found that increasing the distance to a lake by one foot is expected to lead to a 3.1% decrease in the house price (which is equivalent to approximately −10% per one meter increase).

Doss and Taff (1996) examined the value of urban wetlands, including lakes in Ramsey County, Minnesota. Focusing on property with a wetland within a range of 1,000 meters, and considering distance as a linear-continuous variable, they found a large effect of lake view on house values. They discovered its implicit price to be $46,000 for an average property value of $105,000, indicating a value increase of 43.8%; the distance to a lake had smaller effect: an additional 10 meters closer to a lake was worth \$188. A similar large positive waterfront effect (31.7%) was found by Geoghegan et al. (1997) in a region within a 30 mile radius of Washington DC.

Other positive proximity to water effects were found by Mahan et al. (2000) for wetlands in Portland, Oregon. Using a logarithmic transformation of the distance variable, they found relatively smaller effects of proximity to water: distance–price elasticities of 7% for lakes, 1.1% for streams, and 1.8% for wetlands, based on an initial distance of one mile.

Assuming more flexibility by distance levels of the effect of proximity to water on housing value, Orford (2002) studied the effects of proximity to the River Taff in Cardiff, Wales, using distance dummy variables at 50 meter intervals. The results showed that a house which is located within 50 m of the river is expected to be valued nearly £9000 higher (16.8%) than houses in other locations.

It is likely that water bodies in different spatial locations may have dissimilar effects on the prices of housing in their vicinity. In this respect, a Dutch canal or a lake may not be valued similarly to other bodies of water in the US or the UK. However, Luttik (2000) studied the effects of the proximity to water in four locations in the Netherlands and found that the effects of proximity to water on house prices in the Netherlands are relatively similar to the effects measured in other locations. Her results indicate positive price effects of water in residential areas, ranging from 7% for water within 1000 m to 10% for water views.

Cho et al. (2006) also estimated the contribution of water to housing values in multiple places. They specify the house value and distance variables in logarithmic form, and use both standard OLS and locally weighted regressions to estimate the effects of water and green space amenities on housing values in Know County, Tennessee. The global model found a distance to water effect of -2.0%, using an initial distance of one mile. The local model, however, showed regional differences with both positive and negative effects of proximity to water: ranging from -9.0% to $+2.1\%$. They explain that the variation in the results of the local model may result from a variation in the size of the bodies of water and the consequent presence of other positive and attractive amenities in the area.

Anderson and West (2006) conducted a hedonic analysis with US census-block fixed-effects in order to examine the effects of proximity to open-space amenities on housing transaction price. Following estimation of flexible-form models with Box–Cox transformations, they defined the distance variable as a logarithm since they were unable to reject a log-log relation. They find a price–distance elasticity of -0.034 for proximity to a lake, and -0.027 for proximity to a river. Moreover, they also find that omitting local fixed effects and replacing them with a neighborhood control variable causes a bias and in some cases reverses the sign of the estimated effects.

While almost all studies show positive effects of water on housing values, the size of the effects varies greatly, with distance–price elasticities ranging from -9.0% to 43.8%. Moreover, while waterfront locations are particularly valuable, price effects are found to be present up to distances of one mile. Goetgeluk et al. (2005) conclude that the added value of water in residential environments is highly context-dependent. However, although most studies used housing and neighborhood characteristics as control variables in the hedonic analysis, omitted variable bias could still provide an explanation for the differences in estimated results (as was shown in Cho et al. (2006); Anderson and West (2006)). In addition, the studies mentioned above also vary in their definition of the functional relation between proximity to water and housing value. Misspecification of the hedonic price function can seriously undermine its ability to accurately estimate actual willingness-to-pay (Kuminoff et al., 2010; Halvorsen and Pollakowski, 1981; Cheshire and Sheppard, 1995). Often, perhaps inevitably, data availability rather than sound theoretical grounding influences the hedonic model specification (Bryant and Eves, 2013). In this chapter, we address the concern of an omitted variable bias by the use of fixed effects, in a unique spatial setting which allows a very restrictive definition of the fixed-effects groups. Moreover, in order to improve the quality and robustness of the estimated coefficients, we also introduce flexibility in the relation between price and proximity to water. This is addressed by estimating several hedonic specifications in which we reduce restrictions on the functional form of the variables.

5.3 Research methods

5.3.1 Design of the study

Omitted variable bias is often an important concern with hedonic price analyses, since in practice it is hard to make sure that one has controlled for all the relevant characteristics of the houses. For instance, proximity to water makes a site more attractive and therefore bought by people with higher incomes who build more luxury houses. Not all the relevant characteristics may be easily observed, and the result may be that part of the impact of the more luxurious housing is captured in the coefficient for the proximity to water. This effect may also be present in countries like the Netherlands, where residential development is planned. The costs and revenues of such plans are important and the preferences of households for proximity to water and luxury characteristics are, at least to some extent, known by the planners and may be exploited in order to increase the revenues of the project.

Ideally, we would like to be able to compare houses that are identical in all respects except for their proximity to water in order to properly measure the net effect of this characteristic. Fortunately, the Dutch planning system allows us to come very close to this ideal. One consequence of the planned nature of residential development in the Netherlands is that often a number of exactly identical houses are constructed. These houses are exactly equal in floor area, number of rooms, and all kinds of other characteristics. If two such identical houses differ in their proximity to water, this offers an excellent opportunity to measure the value of this amenity.

5.3.2 Fixed-effects model

In this chapter we apply a fixed-effects (FE) approach to estimate the net effect of proximity to water on house prices. The fixed-effects model takes advantage of the panel form of the data in order to control for the fixed characteristics which affect house prices, but remain constant over a similar group of properties. Controlling for these fixed-effects is an important factor in this research since it reduces the suspicion of an estimator bias. In the setup process of the model, we first constructed groups of similar properties which have similar characteristics but differ in their proximity to water. The selection of groups of houses was based on each unique combination of municipality and neighborhood affiliation, year of construction, housing type, and floor area, and is described in more detail in Section 5.4.

In order to examine the effect of proximity to water from other possible angles, we use several specifications of the fixed-effects model. The purpose of this is to increase the flexibility of the estimated values, and to lessen the restrictions on the functional relation between the distance to water and the transaction price.

In the first specification, we use a traditional fixed-effects specification with distance included in the form of a continuous variable:

$$ln(\tilde{P_i}) = \beta_d \tilde{d_i} + \beta_z \tilde{Z_i} + \sum_t^T \beta_t \tilde{M_{t,i}} + \epsilon_i. \tag{5.1}$$

Here, P_i is the transaction price of property i, Z_i is the size of the property, and d_i is its distance in meters from some kind of water. $\sum_t^T \beta_t \tilde{M_{t,i}}$ is a list of year and month dummies which take the value of 1 according to the year and month in which the house transaction occurred, and zero otherwise.

For simplicity, we denote the deviations from the group-mean value of each house i by a tilde: $\tilde{d_i} = d_i - \bar{d_i}$.

A second approach involves estimating the effect of proximity to water in different distance range groups, defined by fixed distance intervals. The specification in (5.2) allows more flexibility for the estimated distance effect coefficients, and it no longer imposes a linearity in the distance for the estimated effect. Each coefficient now captures the effect of proximity to water for groups of houses which are located at 10 meter intervals from the water, by representing these groups with distance dummy variables.

$$
\begin{aligned}
ln(\tilde{P_i}) =&\beta_{10m}I_{10m} + \beta_{20m}I_{20m} + \beta_{30m}I_{30m} + \beta_{40m}I_{40m} + \beta_{50m}I_{50m} + \\
&\beta_z\tilde{Z_i} + \sum_t^T \beta_t\tilde{M}_{t,i} + \epsilon_i
\end{aligned}
\tag{5.2}
$$

where $I_{10m} - I_{50m}$ are indicator variables which are equal to 1 if house i is located within 10–50 meters from a body of water (respectively), and are equal to zero otherwise.

5.3.3 Semi-parametric model specification

The specifications presented above differ in the level of flexibility they allow for the relations between the proximity to water and the house transaction price, but they all impose a functional form on these relations. In order to improve our understanding of the relations between these variable we suggest testing a semi-parametric regression. A semi-parametric regression treats the explanatory variable as an unknown function, and estimates its value without specifying its functional form (5.3):

$$
P_i = f(d_i) + \beta_w W_i + \beta_z Z_i + \sum_t^T \beta_t M_{t,i} + \sum_k \beta_k X_{k,i} + \epsilon_i \tag{5.3}
$$

The semi-parametric specification no longer uses the panel structure of the data. Therefore, in order to maintain consistency in the identification of housing groups, we define dummy variables for each housing group k and include them as parametric variables in the regression, denoted by $X_{k,i}$. In the individual analysis for each of the types of body of water, we define one type to be estimated semi-parametrically while the other type is included as a parametric control variable W_i. In this specification, the transaction price is now defined as an unknown function of the distance to water, $f(d_i)$, and of parametric specifications of the other explanatory variables. The transaction price variable is also no longer defined in its logarithmic values, since the semi-parametric regression ignores functional specifications. Although the results of the semi-parametric regression produce no coefficients for the effect of proximity to water, they can still inform us about the nature of the relation between the variables and point out which is the closest suitable functional form to describe them. In order to construct 95% confidence intervals, during the estimation of the semi-parametric specification we applied bootstrap sampling with 100 replications, which allowed us to compute standard deviations of the estimated coefficients and the smoothed predicted price values (Efron, 1979).

In this research we also aim to make a distinction between the effects of different types of bodies of water on housing types. Therefore, we separate the effect of proximity to a lake from proximity to waterways such as canals or rivers. This distinction is repeated over each of the model specifications mentioned above in (5.1, 5.2 and 5.3).

5.4 Data

5.4.1 Proximity to open water and planned residential development

The Netherlands is a small country in North Western Europe. Ever since the late Middle Ages, it has had a relatively high population density. Water is abundantly present in the Dutch landscape, especially in the western part of the country. Population growth was relatively strong until the 1970s, and the growth in the number of households has been even more pronounced. This gave rise to strong housing demand pressure. To facilitate these developments, the housing stock was increased. To a large extent this happened via the planned construction of new neighborhoods. Usually this meant that a limited number of developers constructed the new housing on the basis of a plan that was agreed upon with the municipality, which played a coordinating, and often also a steering role in the development of the new neighborhoods. These new residential areas were often carefully designed. A mixed housing supply was thought to be desirable to avoid the emergence of areas in which exclusively rich or poor people lived. This called for heterogeneous housing types so as to ensure that affordable housing for different socio-economic groups was available in these areas. On the other hand, it was also advantageous to limit the number of different housing types as this reduced the design costs of these new areas. In the Dutch planning system, terraced housing occupies a prominent place and this reinforced the tendency to construct a relatively large number of identical houses in these new areas.

The Netherlands is also known as the "lowlands" and water (a canal, a river, a pond, or a lake) is often not far away. The presence of open water may also make a site more attractive as a potential new residential area. Moreover, the level of the ground water is high at many sites. This is certainly the case in many new residential areas, and the construction of ponds facilitates the regulation of the ground water level while making the area more attractive as well. Of course, this does not mean that all the houses in these new residential areas are equally close to open water. Within a new residential area, identical houses may easily differ in their proximity to a body of water, for instance because they are both on different ends of a row of houses that starts close to open water. Having open water at a distance of 10 or 50 meters may make an important difference in a household's willingness to pay for this amenity. Houses which are closer to water have a view of the water, and properties adjacent to water may enjoy a direct access to it, and in some cases may also install a small boat dock. On the other hand, it should be noted that houses which are adjacent to water may also be subject to other related negative effects, such as drowning safety hazards or flood risks.

Table 5.1: Number of properties, by proximity to water types

	All water type		Lakes		Waterways	
Distance range (meters)	N	Cumulative percentage	N	Cumulative percentage	N	Cumulative percentage
0 distance	926	5.0	117	0.6	809	4.3
1<10	1,078	10.7	76	1.0	1,007	9.7
11<20	1,446	18.5	118	1.7	1,342	16.9
21<30	1,019	23.9	162	2.5	907	21.7
31<40	1,046	29.5	109	3.1	991	27.1
41<50	1,368	36.8	190	4.1	1,244	33.7
51<60	1,037	42.4	224	5.3	897	38.5
61<70	886	47.1	188	6.3	794	42.8
71<80	806	51.4	182	7.3	726	46.6
81<90	883	56.2	298	8.9	725	50.5
91<100	701	59.9	199	10.0	625	53.9
100 and more	7,495	100.0	16,828	100.0	8,624	100.0
Total	18,691		18,691		18,691	

5.4.2 Selection of the data

The housing transaction data is taken from the database of Kadaster, the Dutch land registry and mapping agency. We selected 113 neighborhoods in 32 large cities in the Netherlands where open water was abundant. The types of water are divided into lakes and waterways, based on the Dutch land registery definitions. Waterways (*Waterloop*) are defined as elongated portions of water in the form of a river, stream, or canal, which are broader than 6 meters. Lakes (*Meer*) are defined as water bodies larger than 50 m^2 that are not a watercourse. This definition may include lakes, ponds, or reservoirs (both natural and artificial). In the neighborhoods that were chosen for this research (Table 5.A.1 in the Appendix provides a list), at least 70% of the total properties have been constructed since 1998. In the selected areas, 18,691 transactions refer to houses that were built in 1998 or later. Moreover, we excluded information about the first sale of the house, since it is likely that identical houses (but not in proximity to water) were sold to the first occupiers at identical prices. It could also be the case that the price paid for the first sale was for the land only, while the price for the construction of the house was paid to a developer. Thus, the selling price of the first sale may have not reflected the complete property value.

Table 5.1 shows that roughly 37% of these transactions are of properties located within 50 meters from some type of water, as measured from the limits of the parcel of each property, and about 60% are located within 100 meters from water.

Our research design requires that we focus on the comparison of identical houses. To do so, we defined homogeneous groups of houses on the basis of the following conditions:

i Located in the same residential area: In the same newly developed neighborhood within a municipality.

ii Identical floor area

iii Identical type, and

iv Identical year of construction.

Although this list of characteristics is short, it is powerful. In particular, the requirement that the number of square meters of living space should match *exactly* is demanding. It is difficult to imagine that an arbitrary pair of houses constructed independently of each other fulfills all these requirements.

Moreover, the maximum distance between two properties in the same fixed-effects group is restricted to 600 meters from the farthest property in the same fixed-effects group.[2]. These relatively short distances ease concerns that within-group differences in neighborhood quality may interfere with the estimated effect of the proximity to water.

[2]Properties which are located further than 600 meters from their FE group (6% of the observations in our data) were excluded from the analysis.

Table 5.2: Descriptive statistics of the housing characteristics

Variable	N	Mean	Std.	Min.	Max.
Year constructed	18,691	2000.59	1.98	1998	2010
Floor size	18,691	135.80	37.07	36	774
Parcel size	18,691	214.56	154.76	32	7,698
Housing price	18,691	271,868	99,669	49,916	975,000

Housing type	N	Percent
Terraced house (*Tussenwoning*)	12,135	64.9
Corner house (*Hoekwoning*)	3,397	18.2
Separate house *(Vrijstaand)*	1,445	7.7
Semi-detached house (*2 onder 1 kap*)	1,417	7.6
Apartment	297	1.6
Total	18,691	100.0

Table 5.2 gives descriptive statistics of the characteristics we use in defining our homogeneous groups as well as about the parcel size. We didn't use the latter variable for defining our groups, because is often happens that parcel sizes differ, for instance because streets are not exactly parallel while the houses on these parcels are exactly identical. We will use parcel size as a control variable in all regressions.

Table 5.3 shows that our selection procedure results in the selection of 7,217 unique groups. Roughly 63% of these groups (24.4% of the observed transactions) have only one observed transaction and they cannot be used in the analysis that follows. In our 'fixed effects' regressions, the variation within groups drives the estimation results. 50% of the transactions belong to groups with at least 5 unique properties. The largest group includes 95 identical properties.

5.5 Estimation results

5.5.1 Models using continuous distance variables

Table 5.4 provides the results of the basic model specifications –distances are included linearly and after logarithmic transformation, per water type, in order to explain the logarithm of transaction price. The table also includes a comparison between the results of the fixed-effects models and the non fixed-effects hedonic regression models. Variables which were used in defining the FE groups, such as floor size, year of construction, housing type, and neighbor-

Table 5.3: Number of homogeneous groups

Group size	N	Number of groups	Percentage	Cumulative percentage
1	4,565	4,565	24.4%	24.4%
2	2,082	1,041	11.1%	35.6%
3	1,476	492	7.9%	43.5%
4	1,176	294	6.3%	49.8%
5	840	168	4.5%	54.2%
6	696	116	3.7%	58.0%
7	735	105	3.9%	61.9%
8	408	51	2.2%	64.1%
9	549	61	2.9%	67.0%
10 to 19	2,960	228	15.8%	82.9%
20 to 29	1,440	60	7.7%	90.6%
30 to 39	539	15	2.9%	93.4%
40 to 49	455	10	2.4%	95.9%
over 50	770	11	4.1%	100.0%
Total	18,691	7,217	100.0%	

hood, are included directly as explanatory variables in the non-FE analysis. Columns 1 and 2 in Table 5.4 compare the results of the non fixed-effects and fixed-effects analyses, where distance is specified linearly. In both estimations we use the natural log of the transaction price as the dependent variable. The effects found in both analyses are relatively small: negative 0.00019 and 0.00009 for the non-FE and FE analysis, respectively, but they are statistically significant and they indicate that there is a considerable difference between the estimated effects. The estimated effect of proximity to water on house prices is 2.1 times higher when fixed-effects are not included. Both estimated effects are much smaller than the values reported in past studies. The coefficient of the FE model implies that every additional meter of distance from a body of water is expected to decrease the housing transaction price by about 0.009%. Although this value is statistically significant, it is extremely small, and its economic significance is almost negligible.

In columns 3 and 4 we distinguish between two types of open water: lakes and waterways. The estimated effect for proximity to a lake is statistically insignificant, but that of proximity to other waterways is significant, and even smaller than the coefficient for all water types (-0.00047). Despite the statistical insignificance of the distance to lake coefficient, testing whether the effects

Table 5.4: Regression results: Distance from water as linear and logarithmic variable

	Dependent variable: Log. transaction price							
	All water	All water	Lakes & waterways	Lakes & waterways	All water	All water	Lakes & waterways	Lakes & waterways
	No FE	FE	No FE	FE	No FE	FE	No FE	FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance to all water	−0.00019***	−0.00009***						
	(0.00004)	(0.00002)						
Distance to lake			−0.000038	−0.000027				
			(0.000028)	(0.000019)				
Distance to waterway			−0.000125**	−0.000047***				
			(0.000052)	(0.000014)				
Distance to all water (log)					−0.0189***	−0.0108***		
					(0.00346)	(0.00269)		
Distance to lake (log)							−0.0125*	−0.0143**
							(0.00657)	(0.00567)
Distance to waterway (log)							−0.0179***	−0.00778***
							(0.00385)	(0.00275)
Parcel Size	0.00046***	0.00091***	0.000465***	0.000908***	0.000459***	0.000898***	0.000460***	0.000898***
	(0.00017)	(0.00015)	(0.000167)	(0.000155)	(0.000166)	(0.000153)	(0.000165)	(0.000154)
Floor size	0.00434***		0.004344***		0.00425***		0.00424***	
	(0.00024)		(0.000249)		(0.000250)		(0.000250)	
Year and month of sale dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of construction dummies	Yes		Yes		Yes		Yes	
Housing type dummies	Yes		Yes		Yes		Yes	
Neighborhood dummies	Yes		Yes		Yes		Yes	
Constant	11.75***	11.81***	11.76***	11.81***	11.78***	11.85***	11.86***	11.92***
	(0.09335)	(0.03613)	(0.092462)	(0.037989)	(0.0941)	(0.0385)	(0.0988)	(0.0479)
Observations	18,691	18,691	18,691	18,691	18,691	18,691	18,691	18,691
R^2	0.83512	0.36141	0.834754	0.360873	0.838	0.363	0.838	0.363
Number of groups		7,217		7,217		7,217		7,217

Notes: (i) Estimations were conducted using clustered standard-errors, by municipality clusters. (ii) Robust standard errors in parentheses.
*p<0.1, **p<0.05, ***p<0.01

of proximity to either of the water types shows that there is no statistically significant difference between the effects.

The results of the regressions in Table 5.4 provide evidence that the effect of proximity to water was biased positively in past studies, and that controlling for unobserved differences in housing and site characteristics strongly reduces the absolute value of the estimated coefficients. However, although the effect found in the non-FE hedonic model (0.019%) is more than double the effect estimated in the FE model, it is still relatively low compared to the effect found in previous studies. One possible explanation for this can be found by considering the chosen study area. The abundance of water and its prevalence in the Dutch landscape could mean that Dutch households attach lower values to proximity to water compared to the value of this amenity in other countries. Nevertheless, even though the effect of water is somewhat smaller in the Netherlands compared to the effect in other study areas, the value of proximity to water is still found to be positively biased by more than 100 percent when fixed-effects are not controlled.

The results of the first specification, which assumes a log-linear relation between the proximity to water and the property transactions price, can be criticized as insufficient for drawing conclusions. The assumption of the model, according to which every additional meter distance from water has the same constant effect on housing prices, may be inappropriate, as it is likely that the effect of the proximity to water weakens after farther distances. Therefore, we estimate the model using a logarithmic specification of the distance to water. A logarithmic specification of the distance to water variable is also useful in order to estimate its elasticity with respect to the housing transaction prices.[3] The results are specified in columns 5 to 8 of Table 5.4.

The result in column 6 shows that the price–distance elasticity obtained by an FE analysis, with respect to all water body types, is estimated to be -0.0108. This implies that an increase by 1% in distance is expected to reduce the housing transaction price by approximately -0.011%. This effect is also much smaller than the elasticity value obtained by the non-FE analysis, shown in column 5 (-0.019%). After making the distinction between the two examined water body types (columns 7,8), the estimated value of the price–

[3]As an additional device to help us determine which functional form is the most appropriate, we also performed a Box–Cox transformation (Box and Cox, 1964) of the model. The results of the Box–Cox estimation show that θ and λ parameter values are small and relatively close to zero. Although still significantly different from zero, the estimated Box–Cox parameters still imply that a linear specification of both the explained and explanatory distance variables is inappropriate, and that a logarithmic specification of both variables should be tested in the model. Results of the Box–Cox estimation are not reported here but are available upon request from the authors.

distance elasticity maintains a somewhat similar pattern. The price–distance elasticity for lakes is approximately -0.014%, and the elasticity for waterways is approximately -0.0078%, both statistically significant. As in the results of column 4, we find that there is no statistically significant difference between the effects of proximity to lakes and waterways. Although the elasticity of lakes is only slightly smaller in the FE analysis than in the non-FE results, comparing FE and non-FE values for waterways we see that the FE-elasticity of waterways is less than one-half of its non-FE value.

These small but statistically significant values are compatible with the results of the previous specification in the sense that the estimated values are statistically significant but much smaller than what was found in previous studies. Furthermore, comparing corresponding values of distances from water and expected prices emphasizes that distance increments in closer proximities may have a stronger effect than those for greater distances.

5.5.2 Using a threshold

One concern with the estimates just described is that they contain many properties that are so far removed from water that no impact should be expected. As noted in Lansford and Jones (1995), if proximity to water has an impact only within a certain distance range, then including properties located at a greater distance to water in the sample will result in a bias towards zero of the estimated coefficient. To take this into account, we estimated the model again, this time assuming that the maximum distance from water is 60 meters. More explicitly, we transformed the distance d_i to a new variable d_i^* is

$$d_i = \begin{cases} d_i & \text{if } d_i \leq 60 \\ 60 & \text{if } d_i > 60. \end{cases}$$

This new distance variable implies that only differences in proximity to water within the interval 0–60 m are taken into account. We should emphasize that this threshold is relatively small compared with the maximum distance of 4,000 feet (roughly 1200 meters) which was defined by Lansford and Jones (1995). We motivate the setting of the threshold at 60 meters first by examining the distribution of properties by distance to water. Since around 45% of the properties in the sample are located within 60 meters of any type of water, the maximum range is still relatively close to water and therefore permits identification of the effect in near proximities. The choice of 60 meters as the threshold distance was also made after experimenting with several other threshold values between 50 and 100 meters. The differences in the estimated coefficients between the thresholds were quite small.[4]

The value of all estimated coefficients for proximity to water have indeed increased, and the price–distance elasticity for any type of water in column 2 of Table 5.5 is 25% higher than in Table 5.4, implying that a 1% increase in distance is expected to reduce housing prices by 0.013%. In column 4 of Table 5.5 we distinguish between lakes and other types of open water, and now we find significant coefficients for waterways but not for lakes. However, the difference between the estimated coefficients of both types of water is again statistically insignificant. The implied price–distance elasticity under a threshold limit for waterways is estimated at −0.0116. This value is only slightly higher than for the elasticity values of around −0.0078 previously estimated in Table 5.4.

Lakes usually have more recreational value than waterways. This is partly reflected in the higher coefficient values (in absolute terms) in Table 5.4 column 8, and in Table 5.5 column 4. However, the effect of proximity to lakes on housing prices shows a lower statistical significance, particularly when the distance limit of 60 meters is Introduced.

One possible explanation for this is that there are relatively few observations within 60 meters from a lake: only 996 properties, or 5.5% of the sample (see Table 5.1). Another explanation may be that the definition of lakes is relatively broad and may include different bodies of water with various sizes, depths, and types of environmental or recreational development (e.g., recreational lakes and natural shallow ponds). The differences in the valuations of such water bodies may cause a bias in the estimated coefficients. [5]

5.5.3 Models using distance dummy variables

In this section we will further explore the nature of the relation between the proximity to water and housing prices, by introducing further flexibility to the fixed-effects regression model. We will now specify the price of housing as a piecewise constant function of the proximity to water. We estimate the

[4]The estimated results for 50–100 m thresholds are given in Appendix 5.C.1.

[5]To control for the effect of the size of the body of water (as in Anderson and West (2006)), we experimented with including the size of the nearest lake variable in logarithms of square meters. However, since most properties in the same fixed-effects groups are in proximity to the same lake (only 663 properties, or approximately 4% of the sample, belong to a group with properties in proximity to different lakes, out of which only 45 properties are located within 100 meters of a lake) the size of the lake is in practice an additional group fixed-effect, and therefore cannot be captured in the fixed-effects estimation for most of the groups. Due to the relatively small number of valid FE observations and low variability in distance to lakes and lake sizes among the properties in which the analysis is valid, we find no statistically significant effect for lake size on housing prices.

Table 5.5: Distance threshold of 60 meters for the distance to water

Dependent variable: Log. transaction price				
	All water No FE (1)	All water FE (2)	Lakes & waterways No FE (3)	Lakes & waterways FE (4)
Distance to all water (log)	−0.0222***	−0.0134***		
	(0.00457)	(0.00427)		
Distance to lake (log)			−0.0250**	−0.0210
			(0.0121)	(0.0146)
Distance to water (log)			−0.0214***	−0.0116***
			(0.00461)	(0.00397)
Parcel Size	0.000459***	0.000898***	0.000460***	0.000899***
	(0.000166)	(0.000154)	(0.000166)	(0.000154)
Floor size	0.00426***		0.00426***	
	(0.000250)		(0.000250)	
Year and month dummies	Yes	Yes	Yes	Yes
Year of construction dummies	Yes		Yes	
Housing type dummies	Yes		Yes	
Neighborhood dummies	Yes		Yes	
Constant	11.79***	11.85***	11.88***	11.93***
	(0.0951)	(0.0407)	(0.103)	(0.0692)
Observations	18,691	18,691	18,691	18,691
R^2	0.837	0.363	0.837	0.363
Number of groups		7,217		7,217

Notes: (i) Estimations were conducted using clustered standard-errors, by municipality clusters. (ii) Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01

Table 5.6: Piecewise constant price function, for properties located less than 60 from water

Dependent variable: Log. transaction price									
		All water No FE (1)		All water FE (2)		Lakes & waterways No FE (3)		Lakes & waterways FE (4)	
All water	10m	0.0731***	(0.0167)	0.0367***	(0.0110)				
	20m	0.0476***	(0.0148)	0.0297***	(0.00655)				
	30m	0.0487***	(0.0125)	0.0313***	(0.00796)				
	40m	0.0310**	(0.0120)	0.0123*	(0.00649)				
	50m	0.000799	(0.00946)	0.00719	(0.00558)				
Lakes	10m					0.0686	(0.0471)	0.0522*	(0.0280)
	20m					0.0692*	(0.0391)	0.0734**	(0.0352)
	30m					0.0400*	(0.0225)	0.0571*	(0.0290)
	40m					0.0243	(0.0274)	0.0394*	(0.0206)
	50m					8.14e-05	(0.0117)	0.0286*	(0.0163)
Waterways	10m					0.0701***	(0.0169)	0.0323***	(0.0114)
	20m					0.0424***	(0.0144)	0.0220***	(0.00451)
	30m					0.0440***	(0.0125)	0.0229***	(0.00615)
	40m					0.0275**	(0.0114)	0.00857	(0.00538)
	50m					−0.000565	(0.0111)	0.00361	(0.00332)
Parcel size		0.000461***	(0.000167)	0.000905***	(0.000151)	0.000462***	(0.000167)	0.000904***	(0.000148)
Floor size		0.00425***	(0.000253)			0.00425***	(0.000254)		
Year and month of sale dummies		Yes		Yes		Yes		Yes	
Year of construction dummies		Yes				Yes			
Housing type dummies		Yes				Yes			
Neighborhood dummies		Yes				Yes			
Constant		11.69***	(0.0935)	11.79***	(0.0344)	11.70***	(0.0997)	11.79***	(0.0337)
Observations		18,691		18,691		18,691		18,691	
R^2		0.838		0.364		0.837		0.365	
Number of groups				7,217				7,217	

Notes: (i) Estimations were conducted using clustered standard-errors, by municipality clusters. (ii) Robust standard errors in parentheses. $^*p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$

impact of proximity to water within each 10 m distance interval in a completely flexible way. Again, comparison between the non-FE analysis and the FE analysis (columns 1 and 2 respectively) shows that the estimated effects are double when no fixed effects are included. The results in column 1 suggest that immediate proximity to water is expected to increase house prices by approximately 7.3%, compared with the premium of 3.67% predicted by the FE model (column 2). This provides further evidence of an upward bias in the estimated coefficients. The results in column 2 of Table 5.6 also show that the effect of proximity to water is generally weakened gradually with distance. The effect of water on housing prices in the nearest 10 m is estimated to be approximately 3.7%, compared with properties located 50 m or further from water. The effect drops in the 10–20 meters range to approximately 2.97%, remains relatively stable at 3% in the 20–30 meters range, then drops to 1.23% between 30–40 meters and then becomes smaller and statistically insignificant in the 40–50 meters range.

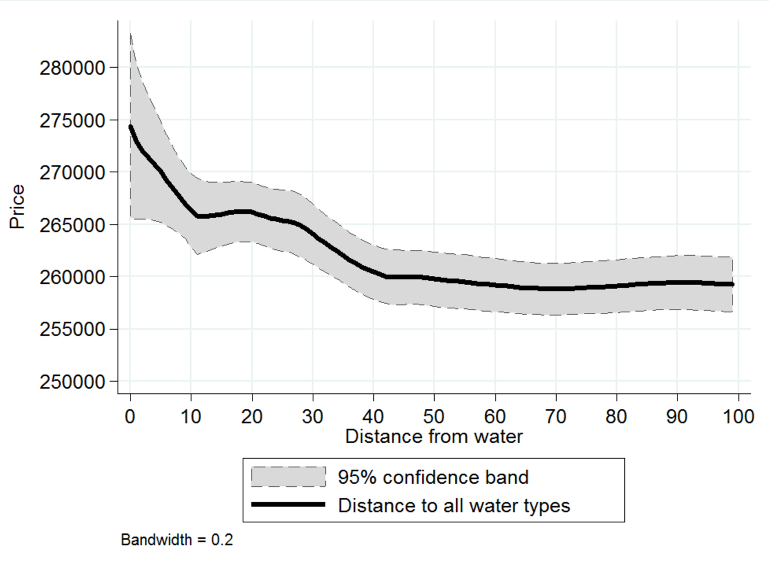
These results also confirm that proximity to water indeed has an extremely local effect, and that the threshold of 60 m which we used above was sufficient to capture the declining effect of distance from water on housing prices. The results for specific types of bodies of water, which are described in columns 3 and 4, support this evidence. The effect of proximity to waterways is statistically significant in closer proximities to water. For the properties which are nearest to waterways, the effect of water is estimated at about 3.2%. As observed in the results for all types of water, the effect drops to 2.2% in the 10–20 meter range, and remains relatively constant at 2.29% in the 20–30 meter range. In the 30–40 meter range, the effect becomes smaller (0.85%) and statistically insignificant.

As in previous model results, The effects of proximity to a lake are generally higher in value, but show a low statistical significance and a somewhat less clear pattern. Considering that only one coefficient is statistically significant at the 5% level (at the 10–20 meter range), and that the estimation is based on a relatively low number of observations, the interpretation of the coefficients is more difficult.

5.5.4 Semi-parametric analysis

In order to get a clearer understanding of the pattern of the effect of water on hosuing prices, we include a semi-parametric model estimation. The use of a semi-parametric model is useful in pointing out which of the discussed functional forms that was tested above is most suitable to describe the relation between the distance from water and the housing price. After removing restrictions on the functional form, the analysis provides a most flexible estimation of the relation between the proximity to water and the house price. The semi-parametric analysis is estimated based on Robinson's double residual estimation (Robinson, 1988), and is specified in Equation 5.3. The results are presented in Figures 5.1, 5.2 and 5.3.

Figure 5.1: Semi-parametric regression of House price on distance from water.



The results in Figure 5.1 show that the effect of proximity to water is strongest in immediate proximities. Consistently with the results of Table 5.6, we see a significant decline in values in the first few meters of distance from water, and then a fluctuating "step-like" pattern in the 20–30 meter range, which suggests that the effects of proximity to water may already have lower impact from this range onwards. The effect of water then gradually declines until it becomes unnoticeable after 75–100 meters.

To check for robustness, and to address the possibility that our estimation may be under or over-smoothed, we also tested whether the decreasing effect of proximity to water is also captured when the semi-parametric analysis is conducted using different bandwidth values.[6] Several techniques have been explored in the literature in order to identify an optimal bandwidth. The most common methods include Silverman's rule of thumb (Silverman,

[6]Bandwidth values correspond with the percentage of the sample used for weighting and predicting smoothed values, where each observation is the center point for weighting, and weights are given following Cleveland (1979) so that zero weights are given to the farthest neighbors according to the predetermined bandwidth value. Stata's `plreg` command, which is used here to estimate the semi-parametric models, uses a bandwidth of 0.8 as default.

Figure 5.2: Semi-parametric regression of House price on distance from waterways.

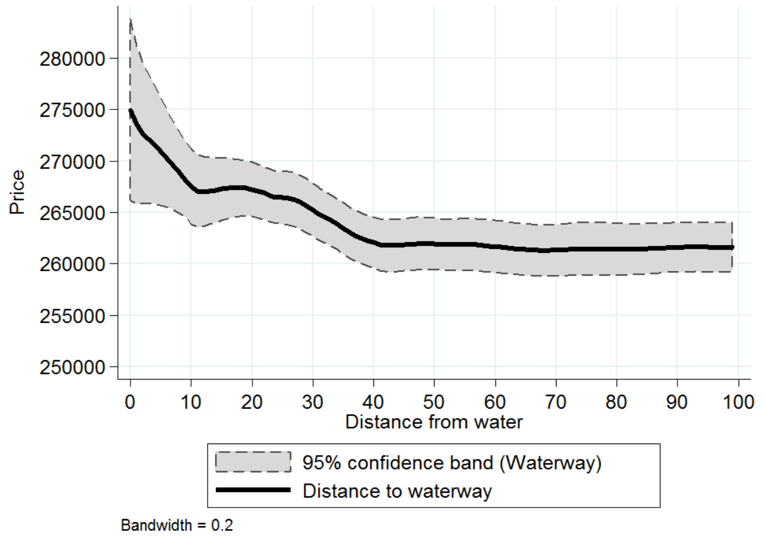
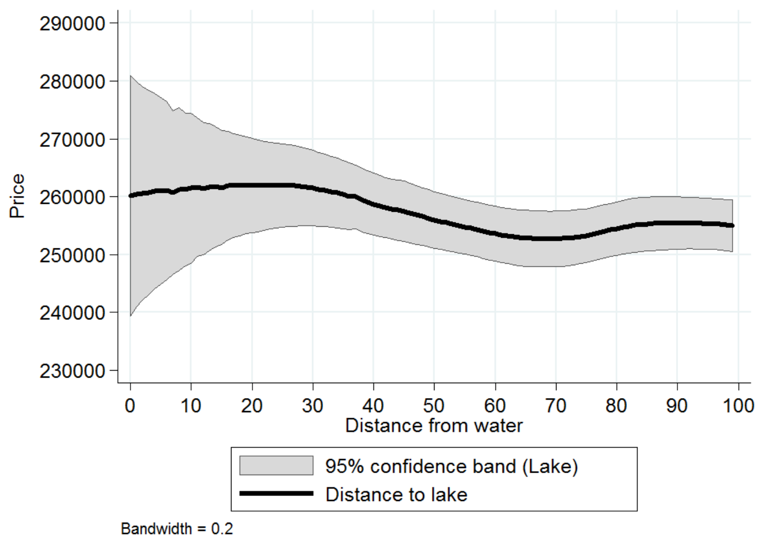


Figure 5.3: Semi-parametric regression of House price on distance from lakes.



1986) or Cross-validation (see discussion in Yatchew (1998)). However, despite the common use of such methods to determine an optimal bandwidth, the estimation may still produce under or over-smoothed values, and a useful indication of the most appropriate smooth parameter can often be obtained by experimentation and visual examination of the results (Yatchew, 1998; Bishop and Timmins, 2008; Koster et al., 2014).[7] Conducting several experiments, we found that a bandwidth of 0.2 is the most appropriate, as it produces smoothed results with relatively small standard errors. Results with different bandwidth values of 0.1, 0.4 and 0.8 are also shown in Appendix 5.D, and they also exhibit a similar trend of decreasing effect with distance of the proximity to water on transaction prices.

The marginal effect of the distance from water on housing prices at the first meter is estimated at roughly 0.54% of the mean housing value[8]. This value can be interpreted as the marginal willingness to pay for residing one meter closer to any type of water (measured at the first meter of distance). For the mean house value, this marginal effect is equivalent to approximately €1,472.

The marginal value of the proximity to all types of water decreases with distance, and becomes almost flat after 10 meters, with 0.19% decline in marginal values, which is equivalent to €516 for the average property. The marginal value of the proximity to water continues afterwards to drop with distance, and is estimated at about 0.029% after 50 meters, or roughly €80. This finding matches the previous parametric models, which also showed a similar steady decline of the effect of water with larger distances, with a slower decline in the 20–30 meter range.

The results of the semi-parametric analysis for waterways (Figure 5.2) also show a similar pattern of decline in the strength of the estimated effect, and the effects become insignificant, and even slightly positive, after approximately 75–100 meters. The marginal willingness to pay at the first meter distance to a waterway is estimated at 0.52% of the housing price, which is roughly €1,425 at the mean value. For 10 meters distance and 50 meters distance, these values drop to, respectively, 0.023% and 0.016% (or €64 and €43).

Despite the fact that positive willingness-to-pay values are observed in several distance ranges, the graphs show a general trend of decline in property value with distance to water. However, the price fluctuations at greater dis-

[7] See the further discussion on the choice of bandwidth in semi-parametric and non-parametric analysis in Cleveland (1979); Fan and Gijbels (1995); Yatchew (1998); Koster et al. (2014).

[8] Housing values are computed based on the predicted smoothed value of the price from the results of the semi-parametric regression, and they reflect the mean predicted value obtained from 100 bootstrap sampling replications.

tances from water show that the effect of proximity to water becomes weaker and less significant with distance. This is particularly apparent when compared with the sharp decline in values in the immediate proximity to water.

As was predicted by the results of the previous models, the semi-parametric analysis produces relatively large confidence bands, which indicates that the effect of the distance to a lake is not significantly different from zero (see Figure 5.3).

5.5.5 Summary of results and marginal effects

The models which were estimated above describe different functional relations between the proximity to water and housing value. Although the majority of the estimated coefficients have been found to be strongly significant and comparable with each other, they correspond to different interpretations of the marginal effects. Table 5.7 provides a summary of these implicit 'premia' or willingness-to-pay values of residing next to water, and attempts compare them. While the distance-dummies model coefficients are directly interpreted as the willingness-to-pay for residing within a particular range of distances from water, the coefficients of the other parametric models provide the marginal effects of an additional meter or percentage distance from water. In order to make them comparable with the distance-dummies, we calculated the premium for residing close to water by adding the effects from the appropriate distance intervals to the 60 meter threshold. The results of the semi-parametric analysis were adjusted similarly, by comparing the mean predicted smoothed value of housing prices in different proximities to water with the predicted house value at the 60 meter threshold.

The comparison shows that all models result in a significant premium for residing directly next to the water (0–1 meters from the water). The log-log model presents a premium value of approximately 6.3%. This predicted value is slightly higher than the prediction of the semi-parametric model, which show that houses in the immediate proximity to water are valued approximately 5.53% more than houses located at the 60 meter threshold or further. With respect to the mean housing value, this percentage is equivalent to about €15,164.

The results from all models show that the willingness-to-pay values decrease with distance. Houses which are located within 10 meters of the water are valued roughly 2%-–4% more than houses with no proximity to water. In addition to the increase in distance from water, the decline between 0–1 meters and 1–10 meters can be explained by the fact that not all houses in the latter range may enjoy a view to the water, and most likely do not have direct (private) access to water, which may affect their willingness-to-pay.

Table 5.7: Willingness to pay values for proximity to water—Model comparison

		0-1 meters		1-10 meters		40-50 meters	
	Model	Percentage of housing value	Implicit value (Euro)	Percentage of housing value	Implicit value (Euro)	Percentage of housing value	Implicit value (Euro)
All water	log-lin	4.41%	11,977.2	3.68%	10,013.7	1.52%	4,123.3
	log-log	6.30%	17,133.0	2.37%	6,447.5	0.56%	1,523.9
	dummies			3.67%	9,969.7	0.72% (ns)	1,953.4 (ns)
	semi-par	5.53%	15,164.6	3.99%	10,785.4	0.28%	715.8
Lakes	log-lin	8.94% (ns)	24,305.7 (ns)	7.47% (ns)	20,321.2 (ns)	3.08% (ns)	8,367.5 (ns)
	log-log	9.84% (ns)	26,759.3 (ns)	3.70% (ns)	10,070.0 (ns)	0.88% (ns)	2,380.2 (ns)
	dummies			5.22% (ns)	14,191.7 (ns)	2.86% (ns)	7,779.1 (ns)
	semi-par	2.51%	6,531.9	2.79%	7,286.8	1.42%	3,663.7
Waterways	log-lin	3.39%	9,216.9	2.83%	7,706.0	1.17%	3,173.0
	log-log	5.47%	14,864.0	2.06%	5,593.6	0.49%	1,322.1
	dummies			3.23%	8,794.7	0.36% (ns)	981.5 (ns)
	semi-par	4.85%	13,325.2	3.43%	9,295.8	0.08%	207.6

Notes: (i) Implicit values in Euro are calculated based on house value of 271,800€, which is the mean house value in the sample. (ii) (ns) indicates that the estimated coefficient is not significant at 5% level. (iii) The reported WTP values for the log-linear and log-log models are interpreted as the implied differences in values of properties in each distance interval, compared with the value of properties which are located at the 60 meter threshold. (iv) Reported values for the distance-dummies model are in reference to the coefficient value for 0-10 meters.

The effects of proximity to water becomes significantly smaller after 40 meters and are hardly apparent. The majority of the models predict that the willingness to pay for the proximity to water in these distance ranges is relatively close to zero.

5.6 Conclusion

In this research we exploited data from new residential developments in a large number of neighborhoods in the Netherlands, and conducted a fixed-effects analysis to explore the relation between the proximity to water and the value of housing. Our restrictive fixed-effects definition creates groups of houses that are identical in terms of housing type, floor size, year of construction, and neighborhood affiliation, and are located in close proximity to each other (often next to each other, as many houses are terraced).

By controlling for fixed effects, we were able to better identify the effects of proximity to water, and to avoid the positive bias which results from the tendency of houses that are located closer to water to be more luxurious, and therefore more expensive. Comparing the fixed-effects model with a non fixed-effects hedonic model, we find that omitted variable bias is responsible for approximately one-half of the estimated effect in the hedonic estimation results. This provides evidence that it is necessary to control for fixed effects in order to reduce the positive estimator bias of the proximity to water.

In addition, we estimated several parametric and semi-parametric model specifications in order to shed light on the patterns in the relation between the proximity to water and housing prices. The estimation results from the specified models and the calculations of the marginal effects and water-premia show that the effect of proximity to water on housing prices is smaller than the values which have been reported in past studies.

The results of the models which were presented raise several issues. First, the effect of proximity to water, as estimated in the specifications above, is smaller than the results from previous studies. We find that proximity to water increases housing value by roughly 4% to 6%, but only within the most immediate proximity to water. This finding strengthens our argument that a tendency to develop higher quality residential housings in close proximity to water may have previously led to a positive bias in the estimated effect of the proximity to water.

Second, we did not find statistically significant differences between the valuations of different types of bodies of water, and the effect of distance from lakes on housing prices is statistically insignificant in most models. Based on the log-linear model, the proximity to lakes seems to be valued slightly higher than the proximity to waterways, and it also decreases slower with distance, but the difference in values is not statistically significant. In addition, the results of the distance dummy model show that while positive effects are estimated in certain distance intervals, it is difficult to identify a clear trend of price decline with distance from a lake. This is also evident from the log-log model and the semi-parametric analysis, in which the effects of the proximity to a lake are not significantly different from zero. The reasons for that are likely to be the relatively small number of observations in proximity to lakes, as well as identification problems of the different types of bodies of water which are categorized together as "lakes."

Third, All models indicate that the effect of proximity to water is extremely local and evident mostly in very close proximity to water. The effect of the proximity to water becomes weaker and gradually decreases with distance until it becomes small and statistically insignificant, more than approximately 60 meters from all types of bodies of water.

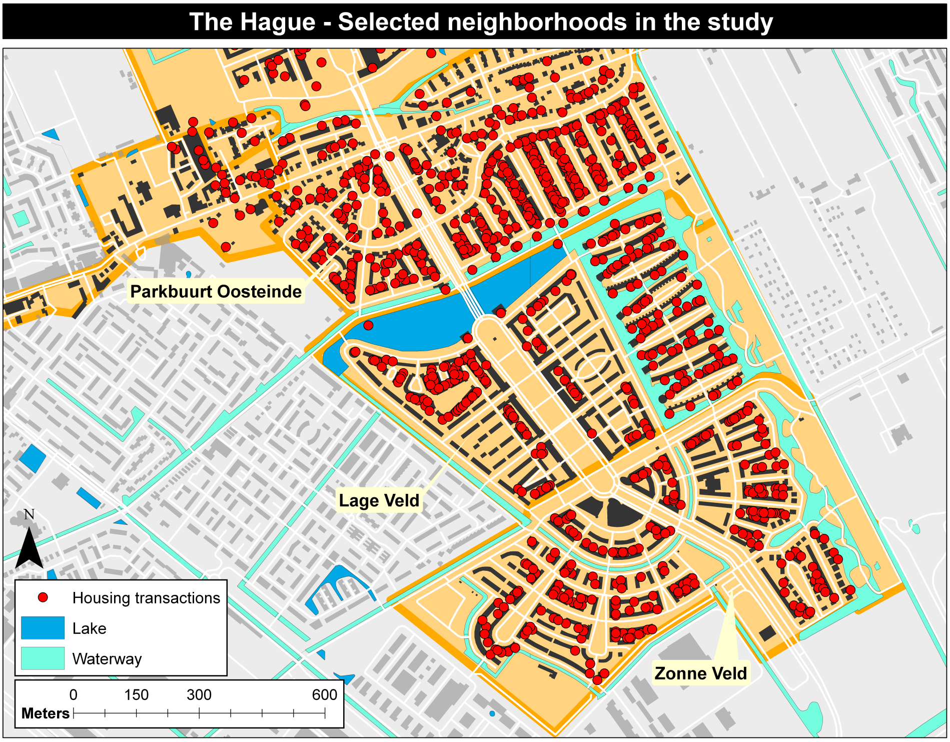
5.A List of neighborhoods in the research

Table 5.A.1: List of Neighborhoods in the research

#	City	Neighbourhood	#	City	Neighbourhood
1	's-Gravenhage	Erasmus Veld	58	Deventer	Spikvoorde
2	's-Gravenhage	Hoge Veld	59	Dordrecht	Dordtse Hout
3	's-Gravenhage	Parkbuurt Oosteinde	60	Dordrecht	Vissersdijk-Oost
4	's-Gravenhage	Lage Veld	61	Dordrecht	De Hoven
5	's-Gravenhage	Zonne Veld	62	Ede	Kernhem
6	's-Gravenhage	Bosweide	63	Eindhoven	Driehoeksbos
7	's-Gravenhage	Morgenweide	64	Eindhoven	Grasrijk
8	's-Gravenhage	Singels	65	Eindhoven	Bos- en Zandrijk
9	's-Gravenhage	Waterbuurt	66	Eindhoven	Bosrijk
10	's-Gravenhage	De Bras	67	Emmen	Delftlanden
11	's-Gravenhage	De Lanen	68	Emmen	Parc Sandur
12	's-Gravenhage	De Velden	69	Enschede	het Brunink
13	's-Gravenhage	De Vissen	70	Groningen	Stadspark
14	's-Hertogenbosch	Broekland	71	Groningen	Dorkwerd
15	's-Hertogenbosch	De Watertuinen	72	Groningen	Bangeweer
16	's-Hertogenbosch	Maasakker	73	Haarlemmermeer	Floriande-West
17	's-Hertogenbosch	Empel-Oost	74	Haarlemmermeer	Floriande-Oost
18	's-Hertogenbosch	De Haverleij	75	Haarlemmermeer	Nieuw-Vennep-Getsewoud-Noord
19	Alkmaar	Vroonermeer-Zuid	76	Haarlemmermeer	Nieuw-Vennep-Getsewoud-Zuid
20	Almelo	Kollenveld-Bolkshoek	77	Helmond	De Veste
21	Almelo	Nijrees	78	Helmond	Schutsboom
22	Almere	De Velden	79	Helmond	Stepekolk
23	Almere	Tussen de Vaarten Noord	80	Hengelo	Het Broek
24	Almere	Tussen de Vaarten Zuid	81	Leeuwarden	Havankpark
25	Almere	Literatuurwijk	82	Leeuwarden	Hemrik
26	Almere	Noorderplassen	83	Leeuwarden	Zuiderburen
27	Almere	Verspreide huizen Almere-Stad	84	Leiden	Roomburg
28	Almere	Oostvaardersbuurt	85	Lelystad	Houtribhoogte-Parkhaven
29	Almere	Seizoenenbuurt	86	Lelystad	Golfresort-Zuigerplasbos
30	Almere	Indischebuurt	87	Lelystad	Landstrekenwijk
31	Almere	Eilandenbuurt	88	Lelystad	Hollandse Hout
32	Almere	Stripheldenbuurt	89	Lelystad	De Landerijen
33	Almere	Sieradenbuurt	90	Lelystad	Flevopoort
34	Amersfoort	Birkhoven en Bokkeduinen	91	Nijmegen	Oosterhout
35	Amersfoort	Centrum	92	Rotterdam	Nesselande
36	Amersfoort	Stadstuin	93	Schiedam	Buurt 98
37	Amersfoort	Waterkwartier	94	Tilburg	Koolhoven
38	Amersfoort	Eindweg en Landweg	95	Tilburg	Witbrant
39	Amersfoort	Velden-Noord	96	Utrecht	Terwijde-West
40	Amersfoort	Velden-Zuid	97	Utrecht	Parkwijk-Noord
41	Amersfoort	Lint-Oost	98	Utrecht	Parkwijk-Zuid
42	Amersfoort	Laak-Zuid	99	Utrecht	Langerak
43	Apeldoorn	Schoonlocht	100	Utrecht	Vleuterweide-West
44	Arnhem	Schuytgraaf-Noord	101	Utrecht	Vleuterweide-Noord/Oost/Centrum
45	Arnhem	Schuytgraaf-Zuid	102	Utrecht	Vleuterweide-Zuid
46	Breda	Waterdonken	103	Utrecht	Veldhuizen
47	Breda	Heilaar	104	Venlo	Hagerbroek
48	Breda	Steenakker	105	Zaanstad	Willis
49	Breda	Kroeten	106	Zaanstad	Waterrijk
50	Delft	Molenbuurt	107	Zaanstad	Parkrijk
51	Delft	Buitenhof-Zuid	108	Zoetermeer	Oosterheem-Zuidwest
52	Delft	Bedrijventerrein Zuideinde	109	Zoetermeer	Oosterheem-Noordoost
53	Delft	Koningsveldbuurt	110	Zwolle	Frankhuis
54	Deventer	Steinvoorde	111	Zwolle	Werkeren
55	Deventer	Graveland	112	Zwolle	Millingen
56	Deventer	Het Jeurlink	113	Zwolle	Holtenbroek I
57	Deventer	Het Fetlaer en Spijkvoorder Enk			

5.B Map of selected neighborhoods

Figure 5.B.1: Map of selected neighborhoods in the study: The Hague *('s Gravenhage / Den-Haag).*



5.C Regression results under different thresholds

Table 5.C.1: Regression results under different thresholds between 50 and 100 meters from water (log-lin and log-log models)

	Dependent variable: Log. transaction price					
Distance threshold (meters)	50m	60m	70m	80m	90m	100m
	(1)	(2)	(3)	(4)	(5)	(6)
Distance to all water (log)	−0.0132***	−0.0134***	−0.0134***	−0.0132***	−0.0131***	−0.0123***
	(0.00422)	(0.00427)	(0.00422)	(0.00419)	(0.00399)	(0.00368)
Parcel Size	0.000898***	0.000898***	0.000898***	0.000898***	0.000898***	0.000899***
	(0.000154)	(0.000154)	(0.000154)	(0.000154)	(0.000154)	(0.000154)
Year and month of sale dummies	Yes	Yes	Yes	Yes	Yes	Yes
Constant	11.85***	11.85***	11.85***	11.85***	11.85***	11.85***
	(0.0406)	(0.0407)	(0.0406)	(0.0405)	(0.0397)	(0.0391)
Observations	18,691	18,691	18,691	18,691	18,691	18,691
R^2	0.362	0.363	0.363	0.363	0.363	0.362
Number of groups	7,217	7,217	7,217	7,217	7,217	7,217

Notes: (i) Regression results shown here are only for distance to all water types. Specific results for distance-to-lake and distance to-waterway are not reported here but available upon request from the authors. (ii) Robust standard errors in parentheses (iii) $p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$

5.D Semi-parametric analysis with different bandwidth values

Figure 5.D.1: Semi-parametric regression of House price on distance from water (Bandwidth = 0.1).

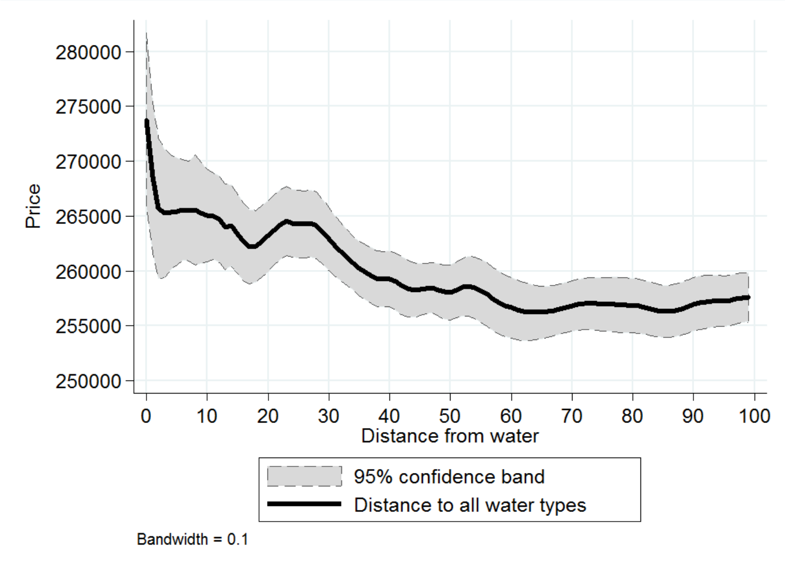


Figure 5.D.2: Semi-parametric regression of House price on distance from water (Bandwidth = 0.4).

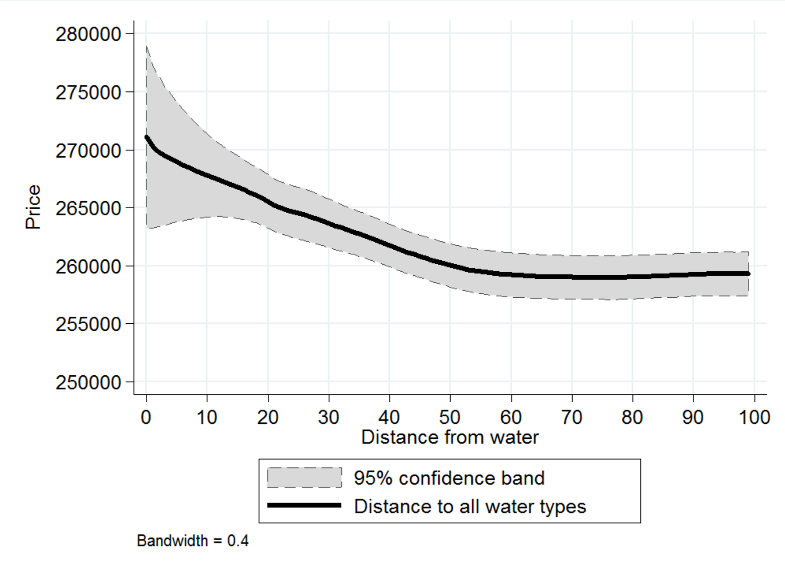
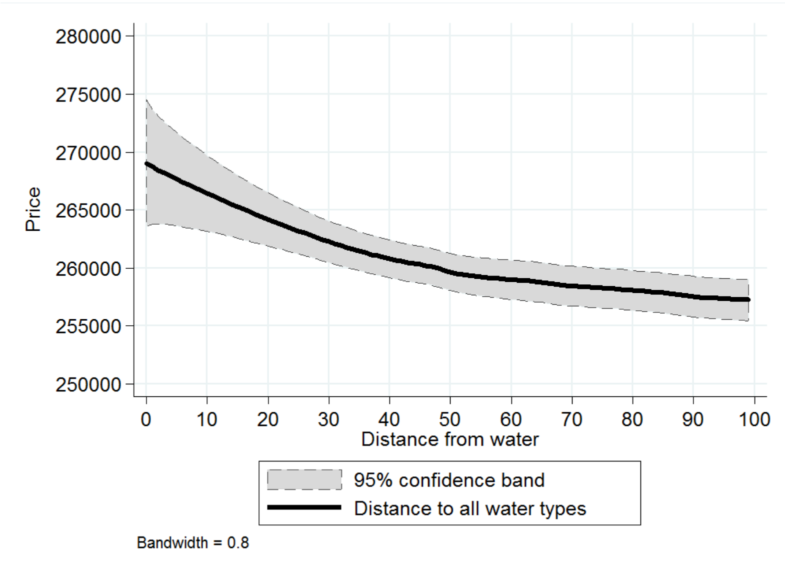


Figure 5.D.3: Semi-parametric regression of House price on distance from water (Bandwidth = 0.8).



Chapter 6

Is the marginal willingness-to-pay for urban amenities different for natives and migrants? Evidence from Amsterdam

6.1 Introduction[1]

Economic growth in urban regions is becoming increasingly dependent on their performance in knowledge-intensive industries (Diamond, 2016; Glaeser and Resseger, 2010; Moretti, 2010; Simon, 1998). Therefore, implementation of urban policies which are aimed at attracting highly educated workers are necessary for cities to grow and compete. Highly educated workers are ever more mobile than before (Artuc et al., 2015), and it is found that by introducing diversity in ideas to the work environment, they increase productivity in innovative industries (Hunt and Gauthier-Loiselle, 2010; Niebuhr, 2010; Suedekum et al., 2014). These trends and findings increase the importance for urban areas to attract these workers, and to consequently raise their competitiveness and potential growth.

The ability of cities to attract highly educated individuals is often related to the attraction power of specific neighborhoods or areas within the city.

[1]This chapter is based on joint work with Jan Rouwendal.
Financial support from the High Education Location Preferences (HELP) research project is gratefully acknowledged.

Metropolitan areas contain a large number of neighborhoods which differ in their attributes, provision of urban amenities and price levels. Since heterogeneous individuals have different location preferences, areas within a city would have different attraction power over population groups. While some neighborhoods would be perceived as very attractive for one group, they could be less appealing for others. For instance, some local amenities may be particularly appreciated by locals, while being hardly valued by foreigners. Knowledge of these valuation patterns can clarify how different parts of the population sort into alternative urban locations, and can also be particularly useful for policy planners who aim to attract specific groups, highly-educated migrants in our particular case.

Ever since Rosen (1974), the standard tool for estimating the value of (urban) amenities is hedonic analysis. However, hedonic analysis offers only limited possibilities for studying differences in the valuation of amenities, which is the focus of interest of the present chapter. Alternative methods are provided by the new economics of equilibrium sorting, reviewed in Kuminoff et al. (2013). These sorting models are structural models of household location choice that explicitly take into account the heterogeneity of the actors and allow estimating the differences in their preferences.

Using such a model, we investigate the amenities which highly-educated individuals of foreign origin find attractive, and compare their preferences with those of local individuals and low- and medium-educated migrants. To do so, we use micro-data from the Dutch population register to estimate a residential sorting model, and analyze the location decisions of individuals in neighborhoods in the Amsterdam area. By examining the residential choices of heterogeneous individuals for diverse neighborhoods we demonstrate the dissimilarity in the values that different individuals attach to various urban amenities. Our results show that locals, low- and medium-educated migrants and highly-educated migrants attach different values to the same amenities, all other characteristics equal. However, highly-educated migrants generally have lower sensitivity to housing prices (controlling for income levels), which is reflected in relatively stronger preference to live in areas richer in urban amenities. We also find that highly-educated migrants also place particularly high values (compared with other groups) towards existing communities of migrants. However, similarly to local individuals and low- and medium-educated migrants, this preference decreases with the increase in the share of migrants. Moreover, highly-educated migrants have the highest marginal willingness-to-pay for areas richer with cultural heritage.

The chapter is organized as follows. Section 6.2 describes the background and reviews previous studies regarding the role of urban amenities as attraction factors, and regarding location choice of individuals, particularly highly-educated foreigners. Section 6.3 describes the data used in the analysis. Section 6.4 explains the methodology and the empirical strategy of the research. Section 6.5 presents the estimation results, implied marginal willingness-to-pay values (MWTP) and discussion, and section 6.6 concludes.

6.2 Urban amenities as attracting factors for human capital

The economic impact of migration on local economies, particularly by inflow of migrants with high human capital, has been the focus of many recent studies. Highly educated workers were found to have a positive effect on regional employment and wages (Behrens and Robert-Nicoud, 2014; Glaeser and Resseger, 2010; Glaeser and Saiz, 2003; Moretti, 2012; Simon, 1998), and foreign highly-educated workers have a particularly positive effect as they increase the stock of regional human capital (Rodríguez-Pose and Vilalta-Bufí, 2005). These findings, along with the increase in mobility of highly educated workers (Artuc et al., 2015; Docquier and Rapoport, 2012), emphasize the potential benefits to regions' economic performance from their ability to attract highly educated workers. Amenities play a significant role in the appealing power of cities, both for local and foreign workers, and they are an important determinant for individuals' location decisions, and consequently also for economic growth of cities (Adamson et al., 2004; Brueckner et al., 1999; Glaeser et al., 2001).

While the role of urban amenities is considered to be central for location choices, some amenities are viewed as more important than others, and the value of similar amenities can often differ between heterogeneous population groups. This is the case for the valuation of housing attributes (Bajari and Benkard, 2005), and also for the valuation of neighborhood characteristics (Bayer et al., 2007, 2004a; Brueckner et al., 1999; Klaiber and Phaneuf, 2010; Van Duijn and Rouwendal, 2013). Differences between groups in valuation of urban amenities instigate a sorting process, in which households sort into urban locations which have the optimal provision of urban amenities in their view (Bayer and Timmins, 2005; Bayer et al., 2004a; Diamond, 2016).

Several previous studies have focused on the preferences of highly-educated local and migrant individuals towards urban amenities, and their central role in determining location choices (Brown and Scott, 2012; Diamond, 2016; Gottlieb and Joseph, 2006). Rodríguez-Pose and Ketterer (2012) find that historical and natural amenities are important in determining regional attractiveness for migrants in European regions. Other studies have also emphasized the importance of existing communities of migrants, as they can lower migration costs and facilitate entry to the labor market through network effects (Åslund, 2005; Damm, 2009; Jaeger, 2007). Additionally, migrant communities can also be viewed as culturally attractive both by locals and foreigners, as they offer a diverse multicultural environment as well as ethnic goods and products (Bakens et al., 2013; Ottaviano and Peri, 2006; Wang et al., 2016). However, an existing community of migrants may also be perceived negatively, as migrants may crowd-out each other as well as other local workers in jobs, and raise social tensions between the local and migrant populations. Saiz and Wachter (2011) document that presence of migrant communities is valued negatively by native households. Examining initial and subsequent location choices of migrants, Bartel (1989) finds that higher-educated migrants to the US from Europe or Asia are likely to relocate away from areas with a large community of migrants of the same origin.

In the Netherlands, the presence of highly-educated workers has also been found to increase employment and productivity of urban areas (Raspe and Van Oort, 2006), and the location choices of such highly-educated workers depend much on the provision of urban amenities in Dutch cities and neighborhoods (Levkovich and Rouwendal, 2014; Van Duijn and Rouwendal, 2013; Van Duijn et al., 2014). While some amenities, like cultural heritage, nature and accessibility, are generally considered attractive by most groups, the value attached to the presence of migrant communities presents more divergent patterns. Bakens et al. (2013) find that cultural diversity, or the presence of migrants from different origins, has a positive effect on wages and housing prices in Dutch cities. However, this effect becomes negative when sorting processes are accounted for, as for some individuals the positive effects of diversity products may be outweighed by negative effects on neighborhood qualities.

Levkovich and Rouwendal (2014) use survey data to estimate a sorting model of locals and migrants for urban amenities in municipalities in the Netherlands, and find that foreign highly-educated individuals place a higher value for cultural heritage and for existing community of migrants, compared with other groups. Here we extend this work by using high quality micro-data from the Dutch population register. Furthermore, we focus on residential choices at the neighborhood level, rather than inter-municipal level, which allows accounting for heterogeneity in characteristics of neighborhoods within cities. Residential location choices between neighborhoods are also more likely to be amenities-oriented, compared with choices between municipalities, where alternatives differ also in their affiliation to different labor markets, and possibly contain many additional unobserved factors. We therefore expect that a neighborhood-level analysis would provide better identification of households' preferences for urban amenities. Although residential sorting model estimations were previously conducted at the urban or neighborhood-level (Bayer et al., 2007; Bayer and McMillan, 2012; Van Duijn et al., 2014), location preferences of migrants were often investigated only at the regional or metropolitan area level. Therefore, this chapter also aims to contribute by providing a richer estimation of the attraction factors of highly-educated migrants, and a comparison of their MWTP values for urban amenities with those of other population groups.

6.3 Data

6.3.1 Study area

The study area concerns 212 neighborhoods in the Amsterdam area, which includes the municipalities of Amsterdam, and its adjacent municipalities of Diemen, Amstelveen, Hoofddorp (Haarlemmermeer), Zaanstad and Ouder-Amstel (see Figure 6.A.1). While the traditionally defined borders of the Amsterdam metropolitan area extend to a larger area (including municipalities at the North Sea coast and most of the provinces of North-Holland and Flevoland), the municipalities which are chosen here as the study area are those that are immediately adjacent to the city of Amsterdam,[2] since they form a single continuous built urban area.

6.3.2 Individual and household data

In order to study the location choices of individuals and households we use micro-data information from the Dutch population register (*GBA*), available from Statistics Netherlands. The micro-data contains the personal characteristics and residential locations of 814,847 individuals over the age of 18 and under the age of 80, residing in the Amsterdam metropolitan area on the 1st of January 2013.[3]

74,023 individuals (approximately 9.1% of the sample) in the data are identified as first-generation migrants to the Netherlands. First generation migrants are identified as individuals who have immigrated since 1999, and both their parents were not born in the Netherlands. Identification of migrants was done using the population register dataset (*GBA*) and by a information from the Dutch Immigration Office (*IND*), both datasets were made available

[2]With the exception of Hoofddorp, which is connected to the rest of the continuous urban area by its direct adjacency to Schiphol international airport.

[3]Approximately 70% of individuals (568,460) are living in the municipality of Amsterdam, 13% (104,260) are living in Zaanstad, and 7% each in Amstelveen (58,400) and Harlemmermeer (57,900).

through Statistics Netherlands. Within the group of migrants, we identify 10,275 highly-educated migrants (1.25% of the sample and 13.8% of migrants). The identification of education level among migrants is based on purpose of arrival, as documented by the Dutch Immigration Office.[4]

Approximately 59.3% of the individuals in the data reside in social housing, provided by housing associations under government-controlled rents that are considerably below market level. Since location decisions of such households largely ignores price consideration,[5] they are excluded from the model estimation.[6]

For the model estimation, we use a randomly selected sample of 131,964 individuals (which consist approximately 16% of the original data) from unique households (no two individuals belong to the same household) who reside in private-sector housing in the Amsterdam area. For these individuals, we observe age, income level, and whether the individual has children, as well as origin and education level (for migrants). Table 6.1 provides a description of the individual characteristics used in the analysis.

6.3.3 Urban amenities

We study the location choice of individuals by examining their preferences towards local housing prices, cultural heritage, natural amenities, distance to

[4]This data identifies highly-educated immigrants from non-EU countries, and relies on their visa status upon arrival to the Netherlands. Immigrants which their purpose of arrival is stated as “study purpose” (*Studie*), “knowledge migrant” (*kennismigrant*), “internship” (*stagiair*) or “research” (*onderzoeker*) are defined as highly-educated. Notably, the data cannot identify the level of education of local highly-educated individuals, not of highly-educated migrants within the EU, who are not obligated to declare their purpose of stay to the Dutch immigration authorities.

[5]Van Ommeren and Van der Vlist (2016) have recently shown that the MWTP of households for social housing can be estimated using information on queuing times for housing availability.

[6]Due to the prevalence of social housing in Amsterdam, discarding such alternatives from the model is likely to have affected our results. This is particularly due to demographic differences which exist between social housing renters and homeowners, particularly in levels of income (in 2013, the average annual gross income of an owner-occupied household was €75,200, compared with €34,700 among renter households (Statistics Netherlands, 2017)). Moreover, eligibility for social housing is restricted based on maximum annual income, while mortgage loans are often approved based on minimal annual income. In addition, the quality of social housing is also generally poorer compared with owner-occupies houses (Van Ommeren and Van der Vlist, 2016). This would imply that marginal willingness-to-pay values for local amenities are expected to be lower for social housing renters. Nevertheless, Van Ommeren and Van der Vlist (2016) show that the MWTP of social renters for house size in the Amsterdam area is on average comparable to that of homeowners, which may imply that discarding renters may not have substantially lowered the estimated MWTP presented here.

Table 6.1: Descriptive statistics of the variables

		Mean	Std.	Min	Max.
Individual characteristics	Age	38.86	11.36	18	80
	Children	0.42	0.49	0	1
	Income	86,337.5	51,282.3	5,000	299,963
	Low-medium educated migrant	0.05	0.22	0	1
	highly-educated migrant	0.01	0.09	0	1
Alternative characteristic	Price	216,179.2	54,811.4	136,558.7	389,383.5
	Share of migrants	4.93	3.93	0	20.0
	Share of migrants (square)	37.49	48.45	0	400
	Mean distance to school	0.71	0.57	0.2	4.0
	Number of monuments	36.69	160.32	0	1,343
	Nature coverage (ha)	101.10	287.28	0	2,431.5
	Mean distance to train station	2.85	1.87	0.40	9.3
	Mean distance to metro station	3.75	3.45	0.24	10.0

Notes: (i) Number of observations of individuals is 131,964 (Sample of individuals residing in the neighborhoods in the study area in 2013). (ii) Number of observations of alternatives is 212 (neighborhoods).

primary schools and the presence of migrants, modeled as the share of migrants in each neighborhood. We also include local accessibility levels to the train and metro networks, which offer inter-urban and suburban public transport accessibility. The housing price in each neighborhood was calculated by a hedonic regression, using housing transactions data from the Dutch Association of Real-Estate agents (*NVM*, 1985–2011). Price of an equal-quality property in our study area varies between €136,500–€389,300 (see Table 6.1 and map in figure 6.B.1).

The share of migrants in each neighborhood was calculated based on the sampled micro-data (Statistics Netherlands). The average share of migrants is approximately 5% (see Table 6.1, Figure 6.1 and map in Figure 6.B.2). 28 neighborhoods have less than 1% migrants, and 15 neighborhoods have more than 10%.[7]

Information on the average distance to primary schools was available from Statistics Netherlands' neighborhood characteristics database (*Wijk en buurtkaart*, Statistics Netherlands, 2011). The average distance to primary schools reflects the density of schools in the neighborhood, and not the quality of the schools in the neighborhood. However, In contrast with the school-districts system in the US, enrollment to primary schools in the Netherlands is not limited by residential location, and parents may enroll their children in school regardless of their choice of residential area. Nevertheless, residential location does matter as waiting lists for high-quality schools are long, and

[7]Only one neighborhood (*Uilenstede* in Amstelveen) has over 15% migrants, as it contains a large university student housing complex.

residential proximity to schools (irrespective of neighborhood boundary) gives priority in enrollment. The quality of schools is often argued to be correlated with unobserved neighborhood characteristics, particularly household income (see in example Bayer and Timmins (2005)). Here we observe the average distance to primary schools in a neighborhood, which does not represent school quality, but may still be correlated with other demographic characteristics, such as the share of households with children. While these arguments for endogeneity are likely to be valid under a school-district enrollment system, the relatively open school enrollment system in the Netherlands may be argued to reduce this suspicion. The average distance to schools indeed demonstrates weak correlation with the average income and the share of households with children.[8] Nevertheless, the average distance to primary schools may still be correlated with other unobserved demographic characteristics. If such characteristics are desirable, the consequence would be over-estimation of the MWTP for schools.

Cultural heritage is measured as the number of official monuments in a neighborhood, as defined by the Cultural Heritage Agency of the Netherlands (RCE). This definition represents the level of cultural heritage in a given area (Van Duijn and Rouwendal, 2013), and is also strongly correlated with other measures of cultural heritage, such as the area covered by buildings constructed before the year 1900 ($\rho = 0.845$), or the year 1800 ($\rho = 0.921$).

Provision of natural amenities and open space is measured using land use spatial information from Statistics Netherlands (2012). We defined open-space as the area in hectares covered by either parks, gardens, forests, natural reserves, wetlands, water bodies, lakes and rivers.

Accessibility variables are included in the analysis to explore the values that households attach to commuting convenience while choosing a residential location. While only 4.9% of daily commuting trips in the Netherlands are by public transport, in large Dutch cities the share is around 10% (Statistics Netherlands, 2012–2014). Choice of commuting mode varies within the Amsterdam area, as public transport is the preferred modal choice of approximately 10% of commuters in the city center, but of over 30% of commuters in the South-Eastern neighborhoods of the city (KiM, 2014). Since this choice is likely related to personal characteristics (such as income, age or children in the household), the variance in its availability across various alternative neighborhoods is likely to affect households' location decision. Neighborhood data on average distance to train stations is available from Statistics Netherlands neighborhoods map (*Wijk en buurtkaart*, Statistics Netherlands 2012).

[8] The correlation of the average distance to schools in a neighborhood with the share of households with children and the average income of is $\rho = 0.10$ and $\rho = -0.10$, respectively.

Distance to metro stations is calculated using GIS, as the average distance for each neighborhood to the nearest metro station (station location data is available from the Dutch land register's BRT map (BRT,*Basisregistratie Topografie*, TOP10NL, 2012).

6.4 Methodology

To investigate the determinants of the residential location choices of individuals, we estimate a sorting model following the approach of Bayer et al. (2004a). The sorting model relies on the core assumption that individuals derive utility from residing in a certain location alternative, and that their utility function depends on the characteristics of each alternative, as well as on their own personal characteristics. Individuals choose a residential location such that it maximizes their utility. The sorting model is based on a multinomial logit (MNL) model, and is used to estimate the choice probabilities of each heterogeneous household to each alternative location. The model is comparable with hedonic models in the sense that it can also be used to reveal tastes and implied marginal willingness-to-pay values for location characteristics, for average households as well as for heterogeneous population groups.[9]

As in other models that extend the MNL model, the sorting model requires attention to the independence of irrelevant alternatives (IIA) property, which is considered as one of the model's main weaknesses. However, allowing for sufficient heterogeneity among individuals in the model (by estimating corresponding heterogeneous coefficients) ensures that while the IIA property still holds at the individual level, it does not hold in the aggregate. In other words, introducing heterogeneous coefficients guarantees that substitution patterns between alternatives would be determined by the data (Berry et al., 1995; Bayer et al., 2004a; McFadden and Train, 2000).

An additional concern in the model is the presence of unobserved heterogeneity between alternatives. Location decisions are likely influenced by unobserved characteristics, and they are also very likely to be particularly correlated with the price of housing, as well as the share of migrants, in each alternative neighborhood. Failing to address these endogeneity issues is likely to result in biased coefficients, choice probabilities and willingness-to-pay values. Following Berry et al. (1995) and Bayer et al. (2004a) we deal with this subject by estimating the model in a two-step procedure, in which the first step is used to estimate heterogeneous coefficients, individuals choice probabilities and the location-specific indirect utility of the mean individual. The second

[9]See outlining discussion in Bayer et al. (2007) on the comparison between hedonic and residential sorting models.

step of the estimation explains the location-specific mean indirect utilities using the examined set of urban amenities and instrumented prices and share of migrants. The price and share of migrants instruments are also constructed following Bayer et al. (2004a), as is explained in detail below.[10]

6.4.1 The residential sorting model – Design of the model

We consider a population of individuals $i = 1 \dots I$ that chooses a residential location out of a given set of alternatives $j = 1 \dots J$. Individual i chooses location j such that it maximizes its indirect utility, based on the provision of urban amenities $k = 1 \dots K$ in each location.

$$U_{i,j} = \sum_{k=1}^{K} \beta_{0,k} X_{k,j} + \sum_{k=1}^{K} \sum_{l=1}^{L} \beta_{k,l} Z_{i,l} X_{k,j} + \xi_j + \epsilon_{i,j} \tag{6.1}$$

where $U_{i,j}$ denotes the indirect utility of individual i from alternative j, $X_{k,j}$ denotes the value of the k-th characteristic of alternative j, $X_{k,j}$ includes all observed location characteristics, among them are also house prices and share of migrants in a given location, and $Z_{i,l}$ denotes the value of the l-th characteristic of household i. The set of $k * l$ coefficients $\beta_{k,l}$ reflects the cross-effect between individual characteristics and urban amenities, or the preferences of individual with characteristic l to location characteristic k.

We express several individual characteristics (age, children dummy and income) as deviations from the mean values $Z_{i,l} = (Z_{i,l} - \overline{Z_l})$, where $\overline{Z_l}$ denotes the sample mean of characteristic l. The use of de-meaned individual characteristics implies that the first term on the right-hand side of (6.1) can be interpreted as the mean utility attached to alternative j, while the second term reflects the deviation from the mean that is associated with individual i's characteristics. Since the dummies for low- and medium-educated and highly-educated migrants are not de-meaned, here, the interpretation of the first term of the right-hand side ($\sum_{k=1}^{K} \beta_{0,k} X_{k,j}$) refers to the utility from each examined amenity derived by the "average" local individual. Furthermore, an additional alternative-specific term ξ_j is added to account for possible unobserved alternative characteristics. The term ξ_j is assumed to be known by the households and to affect their decisions.

[10] While the share of migrants was often instrumented in the literature by historical immigrant gateways to the destination country (Ottaviano and Peri (2006) in the US, Mocetti and Porello (2010) in Italy), such conditions are not applicable in a small country like the Netherlands. Additionally, other instruments which were previously used, such as exploiting a natural experiment of random location assignment as was done by Damm (2012) in Denmark, were also inapplicable in the Netherlands.

We follow Berry et al. (1995) methodology and estimate the model in two steps. We rewrite equation (6.1) as:

$$U_{i,j} = \delta_j + \sum_{k=1}^{K} \sum_{l=1}^{L} \beta_{k,l} Z_{i,l} X_{k,j} + \epsilon_{i,j}, \tag{6.2}$$

with:

$$\delta_j = \sum_{k=1}^{K} \beta_{0,k} X_{k,j} + \xi_j. \tag{6.3}$$

The first step of the sorting model estimation involves estimating equation (6.2). Assuming that individuals maximize their utility by choosing a residential location, choice probabilities can be derived by a logit model estimation.[11] Along with the choice probabilities, the estimation of the logit model provides us with alternative specific constants (ASC's, a constant for each alternative, noted in the model as δ_j), and the interaction coefficients ($\beta_{k,l}$). The probability for each individual to choose a location is then given by:

$$Pr_{i,j} = \frac{e^V_{i,j}}{\sum_{j=1}^{J} e^V_{i,j}} \tag{6.4}$$

where $V_{i,j} = \delta_j + \sum_{k=1}^{K} \sum_{l=1}^{L} \beta_{k,l} Z_{i,l} X_{k,j}$, $j = 1..J$ for any individual i. The cross-effect coefficients and alternative specific constants are then estimated by maximizing the probability that individual i choose location j, using a maximum likelihood procedure.[12] The estimation of the first step largely depends on the equilibrium condition, according to which the demand should equal the supply of houses in each location j. Imposing an equilibrium restriction, we require that the sum of the predicted choice probabilities (over the whole population) would be equal to the existing housing stock in each alternative location (S_j):

$$\sum_{i=1}^{N} Pr_{i,j} = S_j \tag{6.5}$$

Differences between the predicted aggregated probabilities ($\sum_{i=1}^{N} \widehat{Pr_{i,j}}$, expected demand for alternative j) and observed shares are then used for contraction mapping,[13] to solve for δ_j which maximizes the choice probabilities, under the equilibrium condition, as follows:

[11] Assuming that the random term is IID extreme value type I distributed.
[12] For the maximum-likelihood procedure we use the Berndt et al. (1974) algorithm.
[13] Following Berry et al. (1995).

$$\delta_j^{t+1} = \delta_j^t + ln(S_j - \sum_{i=1}^{N} \widehat{Pr_{i,j}}) \tag{6.6}$$

To ensure that the equilibrium condition holds, the estimation of the first step of the sorting model (namely, the alternative specific constants and the cross-effect coefficients) is conducted in an iterative procedure.

The second step of the estimation involves estimating equation (6.3) in which the values of δ_j, (which were estimated in the first step) are explained by a set of observed location characteristics (urban amenities).

6.4.2 Endogeneity in prices and share of migrants

Since prices and the share of migrants are expected to be correlated with the error term ξ_j, the second stage estimation follows a two stage least squares (2SLS) procedure to account for this endogeneity problem. Here, we follow the approach of Bayer et al. (2004a) and generate a price instrument which is computed based on the model and the data. The instrument is defined as the counterfactual prices which would prevail assuming no unobserved heterogeneity between alternatives exists in the model ($\xi_j = 0$)[14]. Once the price instrument is obtained, we subsequently compute the share of migrants instrument by letting individuals sort according to the artificial price levels, and observing the share of migrants which would prevail in each alternative neighborhood, given these artificial equilibrium prices.[15]

The main advantage of this instrumental variable approach is that it exploits the equilibrium properties of the model, and it is computed considering observed exogenous variables and the initially estimated parameters. The artificial price instrument is valid and relevant since it is correlated with the original price levels, but not with the unobserved heterogeneity ξ_j. By letting the share of migrants in each neighborhood adjust according to the new artificial prices, we argue that this counterfactual share of migrants serves as a valid instrument as well, since it conforms with the same validity and relevance conditions as the price instrument.

[14]This follows Bayer et al. (2007). See examples of use of such instruments in Van Duijn and Rouwendal (2013).

[15]Since the instruments rely on the zero unobserved heterogeneity condition, they have to be recalculated after the second stage coefficients are estimated. This is repeated in an iterative process until convergence.

6.4.3 Calculations of the marginal willingness-to-pay for urban attributes

The coefficients estimated in the first and second step of the sorting model can be used to calculate the marginal willingness-to-pay of heterogeneous individuals for the alternative characteristics. The marginal willingness-to-pay for a certain characteristic k is defined as the change in price that would keep individual's utility level constant, or the monetary value that each individual attaches to various urban characteristics. This value depends on the estimated average and individual coefficients, and can be computed by rearranging the model in (6.1). The marginal willingness-to-pay value for characteristic k is derived as:

$$\frac{\partial P_j}{\partial X_{k,j}} = -\frac{\beta_{0,k} + \sum_{l=1}^{L} \beta_{k,l} Z_{i,l}}{\beta_{0,p} + \sum_{l=1}^{L} \beta_{p,l} Z_{i,l}} P_j \tag{6.7}$$

Since individual characteristics are defined to have zero mean, the MWTP of a base-group individual with average characteristic is:

$$\frac{\partial P_j}{\partial X_{k,j}} = -\frac{\beta_{0,k}}{\beta_{0,p}} P_j \tag{6.8}$$

Equations (6.7) and (6.8) emphasize that differences in estimated cross-coefficient values between population groups, particularly with house prices, will lead to differences in MWTP values. These differences in preferences can effectively explain the sorting mechanism of households to alternative locations.

6.4.4 Spatial extensions

The Amsterdam metropolitan area occupies a relatively small geographical area, which consists of densely populated and inter-accessible neighborhoods. This raises the possibility of spatial dependence between alternatives. That is, it may be possible for individuals to reside in one location, but to derive utility from the amenities of the surrounding locations, without incurring high travel costs. To address this, we test for the presence of spatial correlation in the model's residual using Moran's I and Lagrange multiplier test (see Table 6.2).

The results of the Moran I test show statistically significant, but very small, spatial correlation in the residual. The Lagrange multiplier test shows statistically insignificant spatial correlation in the model's residual.[16] These

[16]Lagrange multiplier test for spatial lag shows presence of spatial correlation in the explanatory variables of the model, which is significant at the 5% level. To address this we also estimated the model using additional spatial lag variables. However, the estimation results

Table 6.2: Spatial correlation tests

Spatial Correlation test	Value	p-value
Moran's I	0.0180	0.00000
Langrange multiplier (Spatial error, robust)	0.1033	0.74788

Notes: (i) Tests were computed following Anselin (1988) and Anselin et al. (1996). (ii) Moran's I significance level is computed based on standard normal distribution. Lagrange multiplier significance value are computed based on χ^2 distribution with 1 degree of freedom.

values imply weak spatial correlation in preferences for unobserved amenities. A possible explanation for this can be the presence of spatial barriers between neighborhoods. For instance, many of the Amsterdam area neighborhoods are separated by water bodies such as the rivers Amstel and IJ, or other wide canals. Moreover, artificial barriers such as Schiphol international airport, the port area, rail tracks and the A10 ring road serve as boundaries between several neighborhoods (see figures 6.A.1 and 6.A.2 for illustration). While such barriers do not prohibit flows between neighborhoods, they might still form a clear impediment between some areas.

Nevertheless, as a robustness analysis we conduct an additional estimation of the model using a GMM/IV estimation method (Drukker and Prucha, 2011; Drukker et al., 2013)). The GMM/IV procedure is a two-step estimation procedure which considers the existence of spatial endogenous regressors. It assumes that the error term in (6.3) includes a spatial autoregressive component:

$$\xi_j = \rho W \xi_j + u_j \tag{6.9}$$

where W is a spatial inverse distance matrix and ρ is the spatial autoregressive parameter. The first step of the estimation includes 2SLS estimation of (6.3) assuming no spatial correlation in the residuals, and using the price and share of migrants instruments as specified in 6.4.1. The estimated 2SLS residuals are then used to compute a GMM estimator for $\tilde{\rho}$. In the second step of the procedure we use the estimated spatial autoregressive parameter in a 2SLS estimation of Cochrane-Orcutt transformed model (6.3):

$$\delta_j^* = \sum_{k=1}^{K} \beta_{0,k} X_{k,j}^* + \xi_j \tag{6.10}$$

using spatial lag variables produces statistically insignificant lagged variables coefficients, and therefore is not reported.

where $\delta_j^* = (I - \tilde{\rho}W)\delta_j$, and $X_{k,j}^* = (I - \tilde{\rho}W)X_{k,j}$ The estimated $\beta_{0,k}$ coefficients are then used again to re-estimate the price and migrants instruments, assuming $\xi_j = 0$. The GMM/IV procedure is then iterated and repeated until convergence.

6.5 Results

6.5.1 Results of the first step of the sorting model estimation

Examination of the first step coefficients of the sorting model (see Table 6.3) confirms that individuals with certain personal characteristics have different preferences for certain urban amenities, and that the expected utility from each alternative location is substantially influenced by individual personal attributes. Although the full interpretation of the estimated coefficient values can only be conducted as part of a marginal willingness-to-pay analysis, first-stage estimation results show that personal characteristics have a statistically significant effect on preferences for urban amenities. The difference in valuation between heterogeneous individuals is particularly notable in their sensitivity to housing prices, which largely influences the willingness-to-pay of other amenities, that is expressed in terms of housing prices. For instance, individuals with income higher than average show a lower sensitivity to prices of housing in alternative neighborhoods. Highly-educated migrants are also significantly less sensitive to higher housing prices, compared with local individuals with the same income level, as well as age and parent status. The price sensitivity of low and medium-educated migrants is insignificantly different from that of local individuals.

6.5.2 Results of the second step of the sorting model estimation

The results of the second step estimation are described in Table 6.4. The results of the 2SLS and GMM/IV models exhibit relatively similar coefficients signs, values and levels of statistical significance. Negative and statistically significant price coefficient is expected, as well as a positive and significant coefficient for monuments and mean distance to primary schools, which reflect positive valuation of these attributes by local individuals with mean characteristics. The price coefficient in the 2SLS and GMM/IV models is substantially stronger than the corresponding OLS coefficient. This implies that the price coefficient is positively biased when we neglect to consider the probable correlation with unobserved neighborhood characteristics.

Table 6.3: Results of the first step of the sorting model

Alternative characteristic	Individual characteristic	Coefficient	Standard error
Price	Age	0.02244***	(0.00150)
	Children	−0.8887***	(0.03540)
	log. Income	1.60121***	(0.02613)
	Low-medium educated migrant	0.00003	(0.07514)
	highly-educated migrant	1.40972***	(0.19457)
Share of migrants	Age	−0.0068***	(0.00038)
	Children	−0.1658***	(0.00893)
	log. Income	−0.0798***	(0.00696)
	Low-medium educated migrant	0.32680***	(0.02024)
	highly-educated migrant	0.38993***	(0.05910)
Share of migrants (square)	Age	0.00013***	(0.00002)
	Children	0.00499***	(0.00058)
	log. Income	0.00084***	(0.00045)
	Low-medium educated migrant	−0.0110***	(0.00119)
	highly-educated migrant	−0.0171***	(0.00360)
Mean distance to school	Age	0.02149***	(0.00174)
	Children	−0.2196***	(0.03833)
	log. Income	0.10093***	(0.03096)
	Low-medium educated migrant	0.01570	(0.09170)
	highly-educated migrant	0.40785***	(0.22711)
log. Number of monuments	Age	0.00147***	(0.00017)
	Children	−0.0237***	(0.00430)
	log. Income	−0.0711***	(0.00294)
	Low-medium educated migrant	0.02655***	(0.00856)
	highly-educated migrant	−0.0956***	(0.01867)
Nature coverage	Age	0.00000***	(0.00000)
	Children	0.00021***	(0.00005)
	log. Income	0.00009***	(0.00004)
	Low-medium educated migrant	−0.0000	(0.00014)
	highly-educated migrant	−0.0000	(0.00037)
log. mean distance to train station	Age	0.00714***	(0.00049)
	Children	0.31079***	(0.01119)
	log. Income	0.06433***	(0.00876)
	Low-medium educated migrant	0.04902***	(0.02643)
	highly-educated migrant	−0.0890***	(0.06451)
log. mean distance to metro station	Age	−0.0102***	(0.00037)
	Children	−0.0322***	(0.00850)
	log. Income	−0.0201***	(0.00671)
	Low-medium educated migrant	−0.0424***	(0.01867)
	highly-educated migrant	−0.1538***	(0.04706)
Alternative specific constants		Included	

Notes: (i) Number of observations is 131,964. (ii) Log-likelihood = −593241.3 (iii) 211 alternative specific constants included. (iv) Robust standard errors in parenthesis, *p<0.1, **p<0.05, ***p<0.01

The share of migrants coefficient is positive and statistically significant, and negative and statistically significant when it is specified in square term. This implies that while a share of migrants is generally valued positively, the attraction power of an existing community of migrants declines with its size. The positive coefficient for distance to train stations indicates that residing in proximity to train stations is not viewed as attractive by the mean local individual, possibly due to correlation with undesirable unobserved attributes. The positive but statistically insignificant distance to metro stations implies that we find no evidence of attraction power of metro stations for the mean local individual.

Table 6.4: Results of the second step of the sorting model

	OLS (1)	2SLS (2)	GMM/IV (3)
Price	−1.5196***	−5.1822***	−4.8625***
	(0.79468)	(0.89332)	(1.76940)
Share of migrants	0.67619***	0.51135***	0.50444***
	(0.08317)	(0.09350)	(0.13109)
Share of migrants (square)	−0.0331***	−0.0159***	−0.0151***
	(0.00428)	(0.00482)	(0.01082)
Mean distance to schools	−2.0012***	−2.1255***	−2.1676***
	(0.31818)	(0.35767)	(0.36560)
log. Number of monuments	0.26189***	0.51366***	0.48205***
	(0.10259)	(0.11532)	(0.14752)
Nature coverage	0.00048	0.00067*	0.00062
	(0.00062)	(0.00069)	(0.00069)
log. mean distance to train station	0.66691***	0.91083***	0.92472***
	(0.22022)	(0.24756)	(0.29141)
log. mean distance to metro station	0.42632***	0.17544	0.24797
	(0.17921)	(0.20146)	(0.28732)
Constant	14.4744***	59.2996***	55.3503***
	(9.71519)	(10.9210)	(21.6279)
Price instrument	No	Yes	Yes
Share of migrants instrument	No	Yes	Yes
R^2	0.4599		
Spatial ρ			0.1875

Notes: (i) Number of observations is 212. (ii) Robust standard errors in parenthesis, *p<0.1, **p<0.05, ***p<0.01

6.5.3 Marginal willingness-to-pay for urban amenities

We now use the coefficients estimated in the first and second step of the sorting model in order to calculate the implied marginal willingness-to-pay values for each amenity by each population group (see Table 6.5). Due to the weak evidence of spatial correlation in the error term, we use the 2SLS coefficients for the MWTP calculations.[17]

The mean local individual is willing to pay approximately €89,800 for residing one kilometer closer to a primary school, or €51,200 for 570 meter closer (which corresponds to one standard deviation of mean distance to primary school). The MWTP of low- and medium-educated migrant is similar, but highly-educated migrants are willing to pay approximately 10% more, €99,700 or €56,800 for one kilometer or one standard deviation closer, respectively. These figures are relatively high compared to previous findings (see for instance Black (1999) and Bayer et al. (2007)). A possible explanation for this is that higher abundance of primary schools may still be correlated with additional unobserved amenities or richer neighborhood facilities. Additionally, although we control for the preferences of families with children (who indeed show stronger preferences for residing closer to primary schools) it may also be that closer proximity to schools may also reflect differences in unobserved neighborhood demographics (Bayer et al., 2007).

The presence of monuments and cultural heritage is positively valued by all population groups. highly-educated migrants are willing to pay the most for presence of cultural heritage – €456 for increase in 1% in monuments in the neighborhood (or €16,730, for an increase of one standard deviation in monuments). This value is 11% higher than the willingness-to-pay of locals, which is approximately €408 for 1% increase in monuments (or €14,960 for increase of one standard deviation). The willingness-to-pay of low- and medium-educated migrants for monuments is approximately €429 for 1% increase in monuments (or €15,740 for increase of one standard deviation), but not statistically different from the value attached by locals. The presence of cultural heritage is often associated with complementary economic activities, such as cafés, restaurants and other leisure consumption activities that are commonly attracted by historical scenery. The presence of these related amenities may be partially reflected in the positive MWTP for monuments, which might explain the relatively high value of cultural heritage.

Willingness-to-pay for open space is also highest among highly-educated migrants, at approximately €35.8 for an additional hectare of open-space. This

[17] Since the 2SLS and GMM/IV coefficient values are relatively similar, accounting for presence of spatial correlation in the residual, and using the GMM/IV results for MWTP calculations, produces similar MWTP values to those obtained using the 2SLS results.

is 26% higher than the willingness-to-pay of locals (€28.3), and 36% higher than the willingness-to-pay of low- and medium-educated (€26.3). However, the difference in MWTP between groups is not statistically significant.

Table 6.5: Marginal willingness-to-pay values for urban amenities

	MWTP (Euro)		
variable	Mean (local individual)	Low/medium educated migrant	Highly educated migrant
Schools	−89800.2	−89137.4 †	−99686.9
Monuments	408.24	429.35 †	456.33
Nature	28.31	26.33 †	35.88 †
Distance to train	13668.7	14404.6 †	16940.0
Distance to metro	2084.94 ⋆	1580.02 ⋆	353.237 ⋆

Notes: (i) Values are calculated following equation (6.7) based on the coefficient values estimated in the first and second stages of the sorting model described in equations (6.2, 6.3). (ii) ⋆ indicates that MWTP values were calculated based on second-stage coefficients which are not significant at the 10% level. (iii) † indicates that MWTP values were calculated based on first-stage coefficients which are not significant at the 10% level, but statistically significant second-stage coefficient, which means that MWTP values are not statistically different from those of the base group (local individuals).

Table 6.6: Marginal willingness-to-pay values for share of migrants

	MWTP (thousand Euro)		
Share of migrants	Mean (local individual)	Low/medium educated migrant	Highly educated migrant
0	21.603	35.410	52.305
5	14.860	23.979	33.070
10	8.1181	12.549	13.834
15	1.3757	1.1190	−5.400
20	−5.366	−10.31	−24.63

Note: (i) Values are calculated following equation (6.7) based on the coefficient values estimated in the first and second stages of the sorting model described in equations (6.2, 6.3).

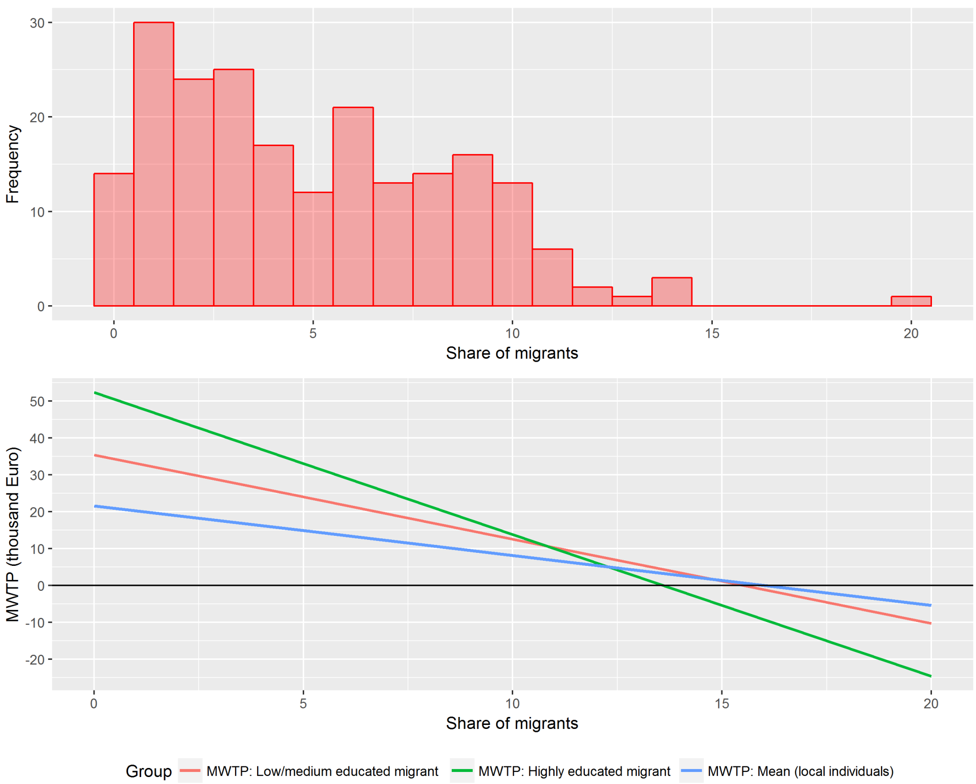
Locating further away from a train station is viewed as more attractive, with a premium of approximately €13,600–€16,900 for every additional kilometer distance from an average neighborhood location. Highly-educated migrants seem to value this most negatively, with MWTP values of €16,940 for every kilometer distance from a train station, 17% higher than low- and medium-educated migrants and 23% higher than locals. MWTP for distance to metro is also positive, indicating negative preferences of all groups towards living in proximity to metro stations. However, a statistically insignificant second-stage coefficient for distance to metro implies that the associated MWTP values and signs cannot be determined. These findings are surprising, and may be explained by a likely correlation between inter-city and suburban public transport accessibility with unobserved unattractive neighborhood characteristics. An additional explanation is that much of the Amsterdam area is connected by the tram and bus systems, which operate more lines and reach more areas compared with the metro network. This suggests that accessibility to public transport does not depend on proximity to accessibility the metro network alone, and therefore its value as a determinant for residential location choice is relatively low.[18]

Our results also show that communities of migrants are valued positively by all groups, when they are relatively small (see Table 6.5 and Figure 6.1), but that the valuation drops with increasing presence of migrants. In low shares of migrants, the presence of migrants is valued substantially higher by migrants of both education level groups. Highly-educated migrants are willing to pay approximately €52,300 for an additional share of migrants, when they are not present. Low- and medium-educated migrants are willing to pay €35,410, and local individuals are willing to pay €21,600 for an additional share of migrants. These values correspond with 33% and 51% lower MWTP compared with highly-educated migrants.

The MWTP values of all groups drop with the increase in share of migrants, but the value that migrants attach to migrant communities drops quicker compared with the value attached by local individuals. The MWTP for migrants become negative after the share of migrants exceeds 13.5% and 15.4% for highly-educated migrants and low- and medium-educated migrants respectively, compared with 16% for local individuals. It is noteworthy that only four neighborhoods in our sample have shares of migrants which are higher than 13.5% (see Figure 6.1), which implies that despite the predictions of the

[18]It may also be argued that the train and metro networks in the Amsterdam area are not important enough to be considered valuable for residential location decisions. However, train and metro separately formed approximately 10% of the total commuting trips in Dutch cities in 2013 (Statistics Netherlands, 2017). While most commuting trips were made by car (43%) and bicycle (30%), train and metro commuting still consisted an important share of such trips.

Figure 6.1: Shares of migrants in Amsterdam neighborhoods, and corresponding marginal willingness-to-pay values



model, and contrary to Saiz and Wachter (2011), the MWTP for migrants for all groups remains positive in almost all cases.

The decline in MWTP value with the share of migrants can be attributed to a subsequent increase in the dominance the negative effects associated with migrant communities, such as possible crowd-out in employment and potential social tensions (Bellini et al., 2008; Ottaviano and Peri, 2006; Ozgen et al., 2011; Suedekum et al., 2014). Both reasons may apply when a sorting process is analyzed on a regional scale. However, since the Amsterdam metropolitan area consists of a single labor market, it is less likely that an increase in the share of migrants in a neighborhood would result in a negative production effect on the employment prospects of migrants (in either education groups). Therefore, the cause of the drop in MWTP values for migrant communities is more likely to be attributed to consumption effects, such as increase in associated social tensions.

6.6 Conclusion

In this chapter we used population register micro-data to analyze the location choices of a large sample of heterogeneous individuals, over neighborhoods in the Amsterdam area. The use of high quality information on individual's personal characteristics, along with an examination of preferences for urban amenities at the neighborhood level, provides reliable estimation of the attraction factors that influence location decisions, and the sorting of heterogeneous household in neighborhoods in Amsterdam.

Our findings show that while preferences towards various urban amenities are generally similar, differences in marginal willingness-to-pay values do exist. Highly-educated migrants are found to attach particularly high values for cultural heritage and schools, and they are willing to pay more for residing in neighborhoods which offer better provision of these amenities. It is also evident that presence of migrant communities influences location decisions of individuals with different attributes. This is particularly eminent in cases where migrants are extremely present, or absent. The attraction power of migrant communities is strongly positive among individuals of all groups, but it declines to potentially negative values when the share of migrants in a neighborhood increases. Given current levels of migrants share in the neighborhoods in our sample, the MWTP values of migrants - particularly highly-educated - are still substantially stronger than those of locals, indicating a much stronger preference for residing where migrants are present.

In general, we find that highly-educated migrants have the strongest preferences for most amenities examined, which results from a lower sensitivity to housing prices (controlling for income level, as estimated in the first stage of the model) and the highest marginal willingness-to-pay values. This could reflect stronger preferences for amenities, but could also possibly be related to poorer information on the local housing market. One primary implication for this finding is that individuals from this group are likely to bid higher prices for attractive neighborhoods, compared with individuals from other heterogeneous groups with equal characteristics, including income level. The outcome of such sorting process is that highly-educated migrants would choose to reside in the neighborhoods which are viewed as most attractive by all groups, given that the share of migrants still remains relatively low. If the supply of housing and the provision of urban amenities remain constant over time, this would also imply that locals and lower educated migrants will be compelled to reside in less attractive neighborhoods in the city, which would likely involve a decline in social welfare. Urban policies which aim to attract highly-educated migrants should therefore focus on maintaining and preserving a variety of amenities – including cultural heritage, open spaces, accessibility to schools and cultural diversity. However, such policies should also consider increasing the supply of housing and provision of urban amenities, to avoid a decline in welfare for local and low- and medium-educated individuals.

6.A Map of the neighborhoods in the study area

Figure 6.A.1: Map of the study area

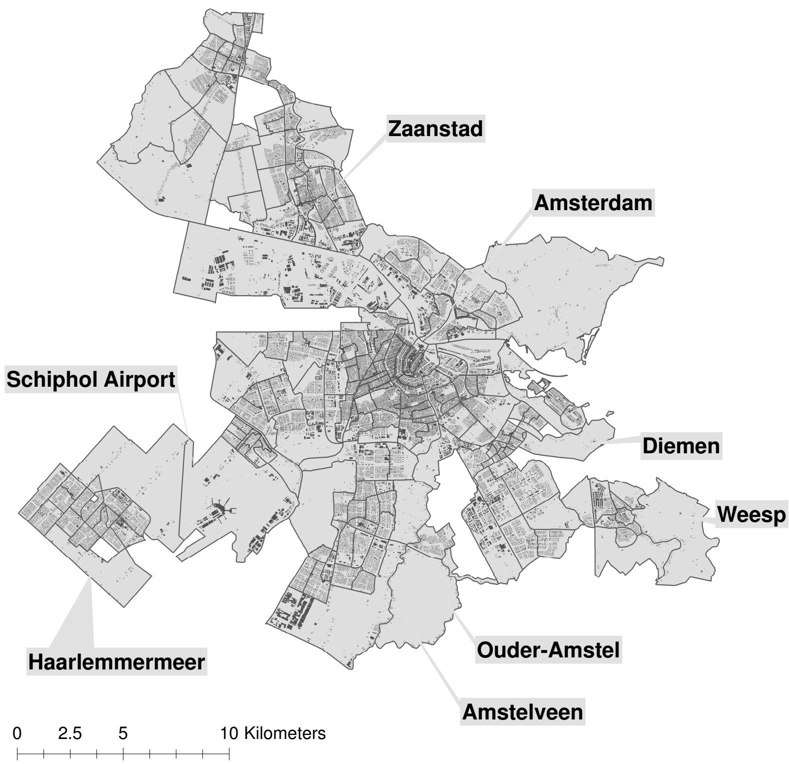
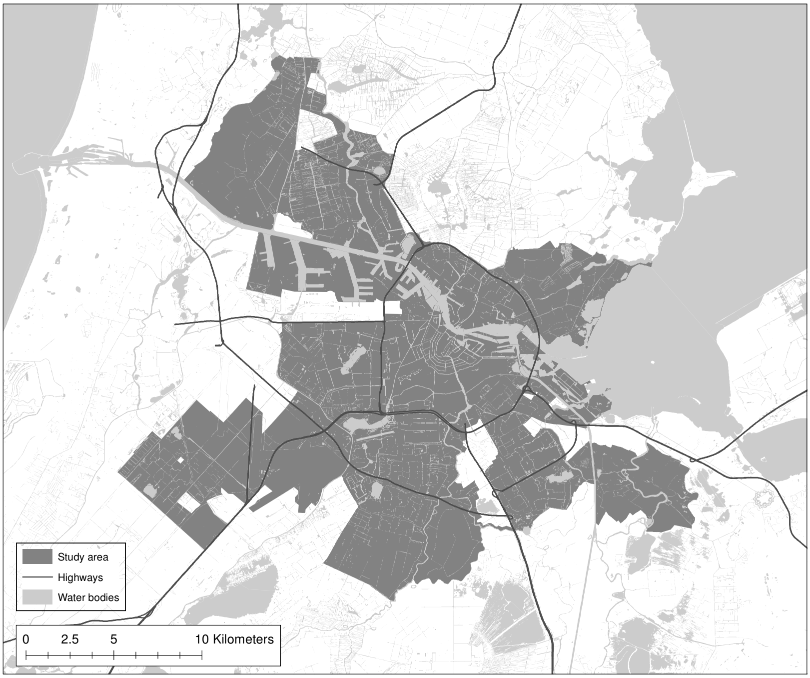
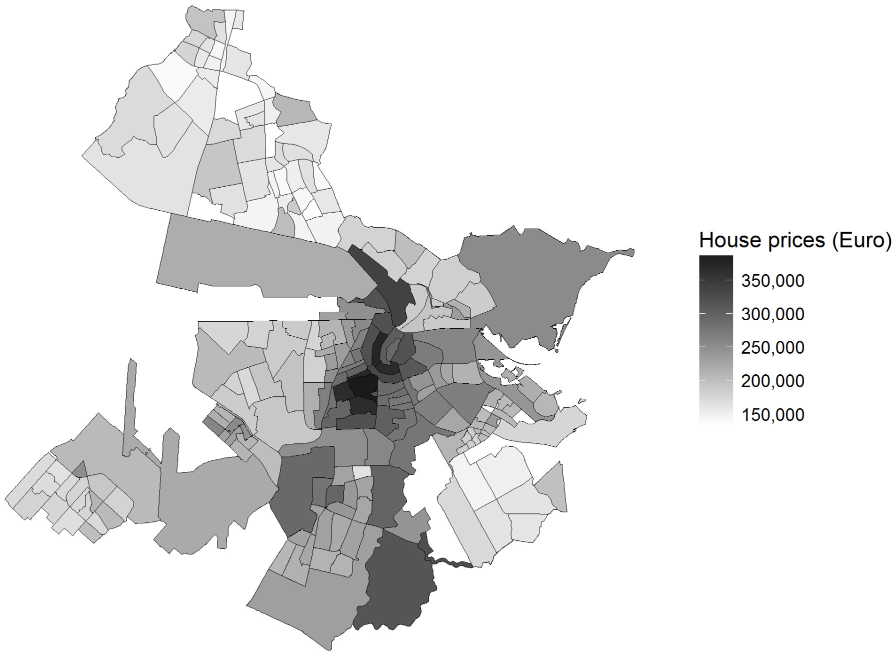


Figure 6.A.2: Water and infrastructure barriers in the study area



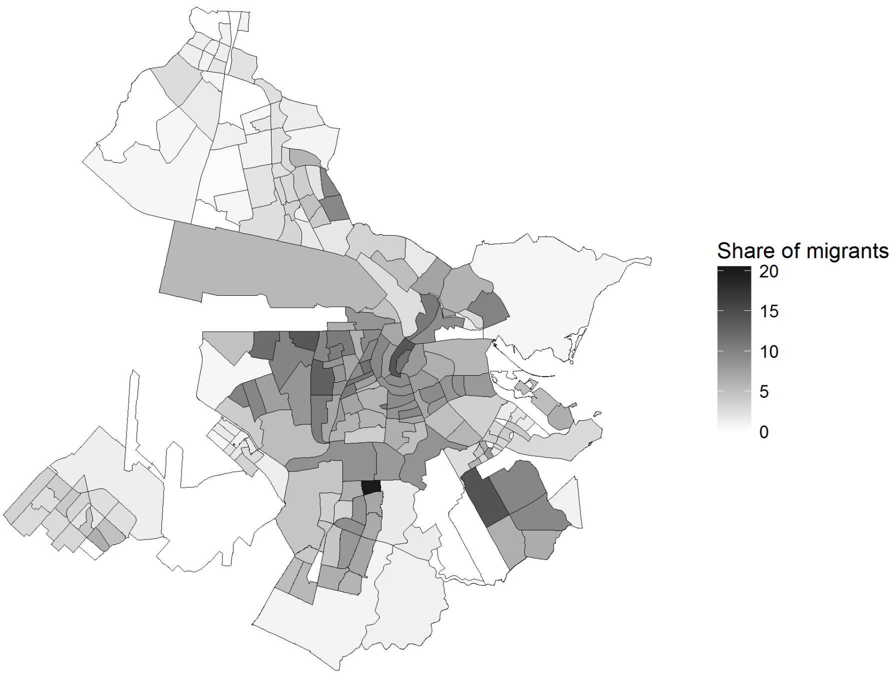
6.B Price levels and distribution of migrants in the Amsterdam area

Figure 6.B.1: House prices in Amsterdam neighborhoods



Calculated using an hedonic regression, based on data from the Dutch real-estate agents association (NVM, 1985–2011).

Figure 6.B.2: Share of migrants in Amsterdam neighborhoods



Migrants are identified as foreign individuals which immigrated to the Netherlands since 1999 (Statistics Netherlands, 2014).

Chapter 7

Conclusion

Urban spatial policies are constantly adjusted and updated in order to accommodate the changing needs of the local population. This often introduces a shock to local housing and land markets, which consequently results in a new market equilibrium. In the course of this dissertation I examined the new equilibrium outcomes of improved accessibility, spatial planning policies and the provision of urban amenities, and demonstrated how they are directly reflected in changes in local spatial patterns and economic conditions.

In chapters 2 and 3, I showed how improved road accessibility, the expansion of the highway network in the Netherlands, has reshaped the landscape and accelerated urban growth in the periphery. Chapter 2 included a repeat-sales analysis of residential transactions in the vicinity of two new highways in the Netherlands. The results showed that new highways are generally valued positively by homeowners and residents, with the exception of locations in close proximity to the highway itself, and that this valuation already becomes apparent several years before the development is completed. Nevertheless, the presence of land development restrictions in the vicinity of large cities had historically hindered the realization of residents' demand for road accessibility. Chapter 3 showed that in contrast to findings in past literature, the expansion of the highway in the Netherlands in the 1960s resulted in a 'leapfrog' of suburbanization to farther remote areas in the periphery. This indicates that strict urban development restrictions in areas adjacent to main cities have severely interfered with the effects of the highway network. The combination of the new highways and the restrictions on urban development has therefore resulted in a large number of small regional centers (part of the "concentrated de-concentration" policy). This has consequently increased commuting times and distance (Schwanen et al., 2001, 2004). These outcomes were found to result in an urban decline in existing cities, and new policies in the following decades have since been aiming to reverse this (Geurs and Van Wee, 2006).

Strict development restrictions are a prominent characteristic of the spatial planning system in the Netherlands. This also extends to a stringent designation of land use, which rarely allows conversion between uses of developed lands. Further development constraints emerge when considering that land designated for residential use is allocated through central national provision, while land designated for commercial use is allocated by local municipal authorities. In practice, this implies that both land uses are subject to different planning regimes. Regulatory obstacles in conversion between land uses, and differences in the provision of the supply of land for development, result in a segmentation of the land market by land use categories. Chapter 4 of the dissertation explored this by an extensive analysis of land transaction data, and showed that undeveloped land parcels designated for residential use are valued almost twice as high as the value of undeveloped industrial land with similar characteristics, and in extreme close proximity. These results imply misallocation of land for development, and that land for residential development is most likely undersupplied.

Urban amenities are often utilized, or artificially developed, in order to increase the appeal of a certain area and attract residents. Relatively high provision of amenities which are perceived as positive can therefore be reflected in the values of nearby real estate properties. This topic is explored in chapter 5, which aimed to capture the value which homeowners and residents attach to living close to water bodies in new residential neighborhoods in the Netherlands. Water is relatively abundant in the Netherlands, and developers of new neighborhoods often exploit water features in order to increase the value of the developed properties. Since houses located on water are viewed as more luxurious by both residents and developers, they may have generally higher quality in other respects, which makes estimating the effect on prices particularly challenging. By examining the effects of proximity to water in model homes, identical in almost all measures except for distance to water, we find that the effect is relatively smaller than in previous findings, and is present only in immediate proximity to water.

The value which individuals attach to urban amenities is not constant, and may vary substantially between individuals with heterogeneous characteristics. Chapter 6 examined how the marginal willingness-to-pay for urban amenities in the Amsterdam area differs for heterogeneous groups, particularly migrants of low- and medium- education and high-education. We investigated observed residential location choices of individuals in the neighborhoods of Amsterdam, and estimated a residential sorting model which reveals how heterogeneous preferences towards urban amenities determine residential location decisions. The findings show that despite differences in marginal willingness-to-pay values, amenities such as cultural heritage, proximity to schools and open space are positively valued by individuals of all groups. The local share of migrants in a neighborhood is also valued positively, but this value decreases with the share of migrants. Highly-educated migrants place a higher value for all examined amenities, indicating a lower sensitivity to prices and a higher willingness to outbid other groups for residing in more attractive areas.

The chapters of this dissertation improve our understanding of how commonly applied urban spatial policies affect local economic conditions, depending on local settings and circumstances. The findings presented here are also strongly linked with possible implications for future policies. First, a better understanding of the economic impacts of various urban policies can primarily be used to guide policies and to minimize possible unwanted indirect effects. One example for this is the relevance for the development of future transportation infrastructure where land development restrictions are present, as was examined in Chapter 3. The findings presented here can support better-informed policy decisions.

Second, the findings presented here may help to avoid inefficient allocation of funding and resources. The examination of the valuation of urban amenities shows that some amenities, such as water bodies, are not a strong attraction factor as was expected by developers, as was shown in Chapter 5. Therefore, the resources dedicated to the development or maintenance of such amenities could be more efficiently allocated elsewhere.

Third, estimated effects on local land and housing prices can be used to quantify the monetary value of a policy impact. This can serve as a base for a second-best compensation mechanism. The latter option can be illustrated by the findings of Chapter 2 regarding the effects of highways on housing prices which vary by proximity to the highway. For instance, possible taxation policy could compensate homeowners in areas in immediate proximity to the new highway, where housing value had decreased, by exploiting value gains from areas in which houses have experienced a value increase in a larger magnitude. Finally, the findings can be used for a wider welfare analysis, which evaluates the full scope of the various policies examined in this dissertation.

Despite efforts to conduct controlled analysis and to identify the effects of spatial policies, the results of the analyses in this dissertation might still be influenced by unobserved features or local trends. Notably, limited availability of regional and historical data on economic indicators and firm activity restricted the analysis from considering equilibrium consequences of urban and spatial policies on the production-side, firm location and local labor markets for instance. It is likely that firms respond to improvements in transportation infrastructure and land development restrictions, as well as to changes in population which are evidently caused by these policy measures. Furthermore, individual location decisions are most likely affected by labor market outcomes (and the associated commuting costs). The ability to account for such factors will provide a more complete picture of urban dynamics. Future research should therefore focus on the examination of policy outcomes, while considering location behavior of firms, and its corresponding reciprocal effect of individual location decisions, commuting patterns and urban structure.

Bibliography

Adamson, D. W., Clark, D. E., and Partridge, M. D. (2004). Do urban agglomeration effects and household amenities have a skill bias? *Journal of Regional Science*, 44(2):201–224.

Agrawal, A., Galasso, A., and Oettl, A. (2017). Roads and Innovation. *Review of Economics and Statistics*, 99(3):417–434.

Ahlfeldt, G. M. and McMillen, D. P. (2014). New Estimates of the Elasticity of Substitution of Land for Capital.

Albouy, D. and Ehrlich, G. (2013). The Distribution of Urban Land Values: Evidence from Market Transactions. *Mimeograph, University of Illinois*.

Alonso, W. (1964). *Location and Land Use. Toward a General Theory of Land Rent*. Harvard University Press, Cambridge, Massachusetts.

Anas, A. and Moses, L. N. (1979). Mode choice, transport structure and urban land use. *Journal of Urban Economics*, 246:228–246.

Anderson, S. T. and West, S. E. (2006). Open space, residential property values, and spatial context. *Regional Science and Urban Economics*, 36(6):773–789.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton university press.

Anselin, L. (1988). *Spatial Econometrics: Models, Methods and Applications*. Kluwer Academic Publishers, Dordrecht.

Anselin, L., Bera, A., Florax, R., and Yoon, M. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*, 26(1):77–104.

Armstrong, R. J. and Rodríguez, D. A. (2006). An Evaluation of the Accessibility Benefits of Commuter Rail in Eastern Massachusetts using Spatial Hedonic Price Functions. *Transportation*, 33(1):21–43.

Artuc, E., Docquier, F., Özden, Ç., and Parsons, C. (2015). A Global Assessment of Human Capital Mobility: the Role of non-OECD Destinations. *World Development*, 65:6–26.

Åslund, O. (2005). Now and forever? Initial and subsequent location choices of immigrants. *Regional Science and Urban Economics*, 35(2):141–165.

Bailey, M. J., Muth, R. F., and Nourse, H. O. (1963). A Regression Method for Real Estate Price Index Construction. *Journal of the American Statistical Association*, 58(304):933.

Bajari, P. and Benkard, C. L. (2005). Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach. *Journal of Political Economy*, 113(6):1239–1276.

Bajari, P. and Kahn, M. E. (2005). Estimating Housing Demand With an Application to Explaining Racial Segregation in Cities. *Journal of Business & Economic Statistics*, 23(1):20–33.

Bakens, J., Mulder, P., and Nijkamp, P. (2013). Economic impacts of cultural diversity in the netherlands: Productivity, utility, and sorting. *Journal of Regional Science*, 53(1):8–36.

Balli, H. O. and Sorensen, B. E. (2013). Interaction Effects in Econometrics. *Empirical Economics*, 45(1):583–603.

Banerjee, A., Duflo, E., and Qian, N. (2012). On the road: Access to transportation infrastructure and economic growth in China. *National Bureau of Economic Research*, (w17897).

Bartel, A. P. (1989). Where Do the New U.S. Immigrants Live? *Journal of Labor Economics*, 7(4):371–391.

Baum-Snow, N. (2007a). Did Highways Cause Suburbanization? *Quarterly Journal of Economics*, 122(2):775–805.

Baum-Snow, N. (2007b). Suburbanization and transportation in the monocentric model. *Journal of Urban Economics*, 62(3):405–423.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., and Zhang, Q. (2017). Roads, Railroads, and Decentralization of Chinese Cities. *The Review of Economics and Statistics*, 99(3):435–448.

Bayer, P., Ferreira, F., and McMillan, R. (2004a). Tiebout Sorting, Social Multipliers and the Demand for School Quality. *National Bureau of Economic Research*, (w10871).

Bayer, P., Ferreira, F. V., and Mcmillan, R. (2007). A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy*, 115(4):588–638.

Bayer, P. and McMillan, R. (2012). Tiebout sorting and neighborhood stratification. *Journal of Public Economics*, 96(11-12):1129–1143.

Bayer, P., Mcmillan, R., and Rueben, K. (2004b). An equilibrium sorting model of sorting in an urban housing market. *National Bureau of Economic Research*, (w10865).

Bayer, P. and Timmins, C. (2005). On the equilibrium properties of locational sorting models. *Journal of Urban Economics*, 57(3):462–477.

Behrens, K., Duranton, G., and Robert-Nicoud, F. (2014). Productive Cities:

Sorting, Selection, and Agglomeration. *Journal of Political Economy*, 122(3):507–553.

Behrens, K. and Robert-Nicoud, F. (2014). Survival of the fittest in cities: Urbanisation and inequality. *Economic Journal*, 124:1371–1400.

Bellini, E., Ottaviano, G., Pinelli, D., and Prarolo, G. (2008). Cultural diversity and economic performance: evidence from European regions. *Hamburg Institute of International Economics (HWWI)*, (3-14).

Berndt, E. K., Hall, B. H., Hall, R. E., and Hausman, J. A. (1974). Estimation and Inference in Nonlinear Structural Models. *Annals of Economic and Social Measurement*, 3(4):653–665.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Bishop, K. and Timmins, C. (2008). Simple, consistent estimation of the marginal willingness to pay function: recovering Rosen's second stage without instrumental variables. *Duke University Department of Economics*.

Black, S. E. (1999). Do Better Schools Matter? Parental Valuation of Elementary Education. *Quarterly Journal of Economics*, 114(2):577–599.

Boeijenga, J., Mensink, J., and Grootens, J. (2008). *Vinex Atlas*. Rotterdam 010.

Box, G. and Cox, D. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B*, 26(2):211–252.

Broitman, D. and Koomen, E. (2015). Residential density change: Densification and urban expansion. *Computers, Environment and Urban Systems*, 54:32–46.

Brown, W. M. and Scott, D. M. (2012). Human Capital Location Choice: Accounting for Amenities and Thick Labor Markets. *Journal of Regional Science*, 52(5):787–808.

Brueckner, J., Thisse, J., and Zenou, Y. (1999). Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. *European Economic Review*, 43:91–107.

Bruinsma, F. and Rietveld, P. (1998). The accessibility of European cities:theoretical framework and comparison of approached. *Environment and Planning A*, 30:499–521.

Bryant, L. and Eves, C. (2013). Data requirements of hedonic models in determining the impact of infrastructure charges on new housing costs. In *RICS Cobra 2013 Conference, Royal Institution of Chartered Surveyors*.

Buitelaar, E. (2010). Window on the Netherlands: Cracks in the Myth: Challenges To Land Policy in the Netherlands. *Tijdschrift voor economische en sociale geografie*, 101(3):349–356.

Buitelaar, E. and Witte, P. (2011). *Financiering van gebiedsontwikkeling een* *empirische analyse van grondexploitaties*. Den Haag, Planbureau voor de Leefomgeving (PBL).

Bun, M. J. and Harrison, T. D. (2014). OLS and IV Estimation of Regression Models Including Endogenous Interaction Terms. *UvA-Econometrics Working Papers*, (2014/02).

Burchfield, M., Overman, H. G., Puga, D., and Turner, M. A. (2006). Causes of sprawl: a portrait from space. *The Quarterly Journal of Economics*, 121(2):587–633.

Capozza, D. R. and Helsley, R. W. (1989). The fundamentals of land prices and urban growth. *Journal of Urban Economics*, 26(3):295–306.

Capozza, D. R. and Helsley, R. W. (1990). The stochastic city. *Journal of Urban Economics*, 28(2):187–203.

Case, K. E. and Shiller, R. J. (1987). Prices of single-family homes since 1970: new indexes for four cities. *New England Economic Review*, (September):45–56.

Chandra, A. and Thompson, E. (2000). Does Public Infrastructure Affect Economic Activity? *Regional Science and Urban Economics*, 30(4):457–490.

Cheshire, P. and Sheppard, S. (1995). On the Price of Land and the Value of Amenities. *Economica*, 62(246):247–267.

Cho, S., Bowker, J., and Park, W. (2006). Measuring the contribution of water and green space amenities to housing values: An application and comparison of spatially weighted hedonic models. *Journal of agricultural and resource economics*, 31(3):485–507.

Cleveland, W. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368):829–836.

Cotteleer, G. and Peerlings, J. H. M. (2011). Spatial planning procedures and property prices: The role of expectations. *Landscape and Urban Planning*, 100(1-2):77–86.

Coulson, N. and Engle, R. (1987). Transportation costs and the rent gradient. *Journal of Urban Economics*, 297:287–297.

Damm, A. P. (2009). Determinants of Recent Immigrants' Location Choices: Quasi-Experimental Evidence. *Journal of Population Economics*, 22(1):145–174.

Damm, A. P. (2012). Neighborhood Quality and Labor Market Outcomes: Evidence from quasi random neighborhood assignment of immigrants.

De Vor, F. and De Groot, H. L. F. (2009). The Impact of Industrial Sites on Residential Property Values: A Hedonic Pricing Analysis from the Netherlands. *Regional Studies*, 35(3):609–623.

Debrezion, G., Pels, E., and Rietveld, P. (2007). The Impact of Railway

Stations on Residential and Commercial Property Value: A Meta-analysis. *The Journal of Real Estate Finance and Economics*, 35(2):161–180.

Dekkers, J. E. C. (2010). *Externalities, land use planning and urban expansion.* PhD thesis.

Diamond, R. (2016). The Determinants and Welfare Implications of US Workers ' Diverging Location Choices by Skill : 1980 – 2000. *The American Economic Review*, 106(3):479–524.

Dieleman, F. M. (1999). The Impact of Housing Policy Changes on Housing Associations: Experiences in the Netherlands. *Housing Studies*, 14(2):251–259.

Dieleman, F. M., Dijst, M. J., and Spit, T. (1999). Planning the Compact City: The Randstad Holland Experience. *European Planning Studies*, 7(5):605–621.

Docquier, F. and Rapoport, H. (2012). Globalization, Brain Drain, and Development. *Journal of Economic Literature*, 50(3):681–730.

Doss, C. and Taff, S. (1996). The influence of wetland type and wetland proximity on residential property values. *Journal of Agricultural and Resource Economics*, 21(1):120–129.

Drukker, D. M., Egger, P., and Prucha, I. R. (2013). On Two-Step Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances and Endogenous Regressors. *Econometric Reviews*, 32(5-6):686–733.

Drukker, D. M. and Prucha, I. R. (2011). A command for estimating spatial-autoregressive models with spatial-autoregressive disturbances and additional endogenous variables. *The Stata Journal*, (1):1–13.

Duranton, G. and Turner, M. A. (2011). The fundamental law of road congestion: Evidence from US cities. *The American Economic Review*, 101(6):2616–2652.

Duranton, G. and Turner, M. A. (2012). Urban Growth and Transportation. *Review of Economic Studies*, 79(4):1407–1440.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26.

Epple, D., Gordon, B., and Sieg, H. (2010). A New Approach to Estimating the Production Function for Housing : Appendices. *American Economic Review*, 100(June):905–924.

Fajgelbaum, P. and Redding, S. J. (2014). External Integration, Structural Transformation and Economic Development: Evidence from Argentina 1870-1914.

Fan, J. and Gijbels, I. (1995). Data-Driven Bandwidth Selection in Local Polynomial Fitting: Variable Bandwidth and Spatial Adaptation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(2):371–394.

Fischel, W. A. (2001). *The homevoter hypothesis: How home values influence local government taxation, school finance, and land-use policies.* Harvard University Press, Cambridge, MA.

Franklin, J. and Waddell, P. (2003). A hedonic regression of home prices in King County, Washington, using activity-specific accessibility measures. *Board 82nd Annual Meeting, Washington.*

Garcia-López, M.-À. (2012). Urban Spatial Structure, Suburbanization and Transportation in Barcelona. *Journal of Urban Economics*, 72(2-3):176–190.

Garcia-López, M.-À., Holl, A., and Viladecans-Marsal, E. (2015). Suburbanization and highways in Spain when the Romans and the Bourbons still shape its cities. *Journal of Urban Economics*, 85:52–67.

Garcia-López, M.-À., Solé-Ollé, A., and Viladecans-Marsal, E. (2014). Do Land Use Policies Follow Road Construction. *CESifo working papers*, (4672).

Geoghegan, J., Wainger, L. a., and Bockstael, N. E. (1997). Spatial landscape indices in a hedonic framework: An ecological economics analysis using GIS. *Ecological Economics*, 23(3):251–264.

Geurs, K. T. (2006). Job accessibility impacts of intensive and multiple land-use scenarios for the Netherlands' Randstad Area. *Journal of Housing and the Built Environment*, 21(1):51–67.

Geurs, K. T. and Van Wee, B. (2006). Ex-post Evaluation of Thirty Years of Compact Urban Development in the Netherlands. *Urban Studies*, 43(1):139–160.

Glaeser, E. L., Gyourko, J., and Saks, R. (2006). Urban growth and housing supply. *Journal of Economic Geography*, 6(1):71–89.

Glaeser, E. L. and Kahn, M. E. (2004). Sprawl and Urban Growth. *Population English Edition*, 4(04):2481–2527.

Glaeser, E. L., Kolko, J., and Saiz, A. (2001). Consumer city. *Journal of Economic Geography*, 1(1):27–50.

Glaeser, E. L. and Resseger, M. G. (2010). The complementarity between cities and skills. *Journal of Regional Science*, 50(1):221–244.

Glaeser, E. L. and Saiz, A. (2003). The Rise of the Skilled City. *National Bureau of Economic Research*, (w10191).

Glaeser, E. L. and Ward, B. A. (2009). The causes and consequences of land use regulation: Evidence from Greater Boston. *Journal of Urban Economics*, 65(3):265–278.

Goetgeluk, R., Kauko, T., and Priemus, H. (2005). Can Red Pay for Blue? Methods to Estimate the Added Value of Water in Residential Environments. *Journal of Environmental Planning and Management*, 48(1):103–120.

Gottlieb, P. D. and Joseph, G. (2006). College-To-Work Migration of Technology Graduates and Holders of Doctorates Within the United States. *Journal of Regional Science*, 46(4):627–659.

Green, R. K., Malpezzi, S., and Mayo, S. K. (2005). Metropolitan-specific estimates of the price elasticity of supply of housing, and their sources. *American Economic Review*, 95(2):334–339.

Gutiérrez, J. (2001). Location, economic potential and daily accessibility: an analysis of the accessibility impact of the high-speed line Madrid–Barcelona–French border. *Journal of transport geography*, 9(4):229–242.

Gutiérrez, J., Condeço-Melhorado, A., and Martín, J. C. (2010). Using accessibility indicators and GIS to assess spatial spillovers of transport infrastructure investment. *Journal of Transport Geography*, 18(1):141–152.

Hahn, J., Ham, J. C., and Moon, H. R. (2011). The Hausman Test and Weak Instruments. *Journal of Econometrics*, 160(2):289–299.

Hajer, M. and Zonneveld, W. (2000). Spatial Planning in the Network Society-Rethinking the Principles of Planning in the Netherlands. *European Planning Studies*, 8(3):337–355.

Halvorsen, R. and Pollakowski, H. (1981). Choice of functional form for hedonic price equations. *Journal of Urban Economics*, 49:37–49.

Hamilton, B. (1978). Zoning and the exercise of monopoly power. *Journal of Urban Economics*, pages 116–130.

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica*, 46(6):1251.

Henneberry, J. (1998). Transport investment and house prices. *Journal of Property Valuation and Investment*, 16(2):144–158.

Hilber, C. A. L. and Robert-Nicoud, F. (2013). On the origins of land use regulations: Theory and evidence from US metro areas. *Journal of Urban Economics*, 75(1):29–43.

Hilber, C. A. L. and Vermeulen, W. (2016). The Impact of Supply Constraints on House Prices in England. *Economic Journal*, 126(591):358–405.

Hsieh, C. T. and Moretti, E. (2017). Housing Constraints and Spatial Misallocation. *National Bureau of Economic Research*, (w21154).

Huang, H. and Tang, Y. (2012). Residential land use regulation and the US housing price cycle between 2000 and 2009. *Journal of Urban Economics*, 71(1):93–99.

Hunt, J. and Gauthier-Loiselle, M. (2010). How Much Does Immigration Boost Innovation? *American Economic Journal: Macroeconomics*, 2(2):31–56.

Iacono, M. and Levinson, D. (2012). Location, Regional Accessibility, and Price Effects. *Transportation Research Record: Journal of the Transporta-*

tion Research Board, 2245(1):87–94.

Ihlanfeldt, K. R. (2007). The effect of land use regulation on housing and land prices. *Journal of Urban Economics*, 61(3):420–435.

Imbens, G. W. and Wooldridge, J. M. (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, 47(1):5–86.

Jaeger, D. A. (2007). Green Cards and the Location Choices of Immigrants in the United States, 1971-2000. *IZA Discussion Papers*, (2145).

Jobse, R., Kruythoff, H., and Musterd, S. (1991). Groeikernen in een veranderend ruimtelijk en demografisch krachtenveld. In *Geografisch Tijdschrift*, volume 25, pages 148–160.

Jun, M. J. and Hur, J. W. (2001). Commuting costs of "leap-frog" new town development in Seoul. *Cities*, 18(3):151–158.

KiM (2014). Mobility Report. Technical report, KiM Netherlands Institute for Transport Policy Analysis.

Kim, K. S., Park, S. J., and Kweon, Y.-J. (2007). Highway traffic noise effects on land price in an urban area. *Transportation Research Part D: Transport and Environment*, 12(4):275–280.

Klaiber, H. A. and Phaneuf, D. J. (2010). Valuing Open Space in a Residential Sorting Model of the Twin Cities. *Journal of Environmental Economics and Management*, 60(2):57–77.

Klaiber, H. A. and Smith, V. (2010). Valuing Incremental Highway Capacity in a Network. *National Bureau of Economic Research*, (w15989).

Kok, N., Monkkonen, P., and Quigley, J. M. (2014). Land use regulations and the value of land and housing: An intra-metropolitan analysis. *Journal of Urban Economics*, 81:136–148.

Koomen, E. and Dekkers, J. (2013). The Impact of Land-Use Policy on Urban Fringe Dynamics: Dutch Evidence and Prospects. In Malkinson, D., Benenson, I., and Czamanski, D., editors, *Modeling of Land-Use and Ecological Dynamics.*, pages 9–35. Springer, Berlin Heidelberg.

Koomen, E., Dekkers, J., and Van Dijk, T. (2008). Open-space preservation in the Netherlands: Planning, practice and prospects. *Land Use Policy*, 25(3):361–377.

Koopmans, C., Rietveld, P., and Huijg, A. (2012). An Accessibility Approach to Railways and Municipal Population Growth, 1840–1930. *Journal of Transport Geography*, 25:98–104.

Koster, H. R. A., Van Ommeren, J. N., and Rietveld, P. (2010). Estimating the Benefits of Improved Rail Access. *Tinbergen Institute Discussion Paper*, (094/3).

Koster, H. R. A., Van Ommeren, J. N., and Rietveld, P. (2014). Agglomeration economies and productivity: A structural estimation approach using commercial rents. *Economica*, 81(321):63–85.

Kramer, H. (2005). Dataset Historisch Grondgebruik Nederland [HGN (1960)].

Kruythoff, H. and Teule, R. (1997). Vinex policy moves into the implementation phase. *Netherlands Journal of Housing and the Built Environment*, 12(1):113–133.

Kuiper, R. and De Regt, W. J. (2007). Landbouw in de Nationale Landschappen. *Netherlands Environmental Assessment Agency (PBL)*, (500074004).

Kuminoff, N. V., Parmeter, C. F., and Pope, J. C. (2010). Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities? *Journal of Environmental Economics and Management*, 60(3):145–160.

Kuminoff, N. V., Smith, V. K., and Timmins, C. (2013). The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets. *Journal of Economic Literature*, 51(4):1007–1062.

Lansford, N. J. and Jones, L. (1995). Recreational and aesthetic value of water using hedonic price analysis. *Journal of Agricultural and Resource Economics*, 20(2):341–355.

Levkovich, O. and Rouwendal, J. (2014). Location choices of highly educated foreign workers: the importance of urban amenities. *Tinbergen Institute Discussion Paper*, (093/VIII).

Ligtermoet, D. M. (1990). Beleid en planning in de wegenbouw : de relatie tussen beleidsvorming en planning in de geschiedenis van de aanleg en verbetering van rijkswegen. Technical report, Ministerie van Verkeer en Waterstaat.

LNV (2008). Plattelandsontwikkelings programma 2007-2013 voor Nederland (POP2). Technical report, LNV (Ministerie van Landbouw, Natuur en Voedselkwaliteit); VROM (Ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieu);Ministerie van Verkeer en Waterstaat.

López, E., Gutiérrez, J., and Gómez, G. (2008). Measuring Regional Cohesion Effects of Large-scale Transport Infrastructure Investments: An Accessibility Approach. *European Planning Studies*, 16(2):277–301.

Lörzing, H., Klemm, W., van Leeuwen, M., and Soekimin, S. (2006). VINEX! Een morfologische verkenning. Technical report, Ruimtelijke Planbureau, The Hague.

Louw, E. and Bontekoning, Y. (2007). Planning of industrial land inf the Netherlands: Its rationales and consequences. *Tijdschrift voor Economische en Sociale Geografie*, 98(1):121–129.

Louw, E., Van der Krabben, E., and Priemus, H. (2003). Spatial development policy: Changing roles for local and regional authorities in the Netherlands.

Land Use Policy, 20:357–366.

Luttik, J. (2000). The value of trees , water and open space as reflected by house prices in the Netherlands. *Landscape and Urban Planning*, 48:161–167.

Mahan, B. L., Polasky, S., and Adams, R. M. (2000). Valuing urban wetlands: A property pricing approach. *Land Economics*, 76(1):100–113.

Martínez, L. and Viegas, J. (2009). Effects of Transportation Accessibility on Residential Property Values. *Transportation Research Record: Journal of the Transportation Research Board*, 2115(1):127–137.

Mayer, C. J. and Somerville, C. (2000). Land use regulation and new construction. *Regional Science and Urban Economics*, 30(6):639–662.

McFadden, D. and Train, K. (2000). Mixed MNL Models For Discrete Response. *Journal of Applied Econometrics*, 15(5):447–470.

Michaels, G. (2008). The Effect of Trade on the Demand for Skill - Evidence from the Interstate Highway System. *The Review of Economics and Statistics*, 90(4):683–701.

Mills, E. S. (1967). An Aggregative Model of Resource Allocation in a Metropolitan Area. *The American Economic Review*, 57:197–210.

MNP (2005). Natuurbalans 2005. Technical report, MNP (Milieu- en Natuurplanbureau), report n. 408763002.

Mocetti, S. and Porello, C. (2010). How does immigration affect native internal mobility? New evidence from Italy. *Regional Science and Urban Economics*, 40(6):427–439.

Mohring, H. (1965). Urban highway investments. In *Measuring benefits of government investments*, pages 231–291. Brookings, Washington, DC.

Moretti, E. (2010). Local Multipliers. *American Economic Review*, 100(2):373–377.

Moretti, E. (2012). *The New Geography of Jobs*. Houghton Mifflin Harcourt.

Muth, R. F. (1969). *Cities and housing: the spatial pattern of urban residential land use*. University of Chicago Press.

Needham, B. (1997). Land policy in the Netherlands. *Tijdschrift voor economische en sociale geografie*, 88(3):291–296.

Needham, B. and Louw, E. (2006). Institutional Economics and Policies for Changing Land Markets: The Case of Industrial Estates in the Netherlands. *Journal of Property Research*, 23(1):75–90.

Nelson, J. P. (1982). Highway noise and property values: a survey of recent evidence. *Journal of Transport Economics and Policy*, 16(2):117–138.

Niebuhr, A. (2010). Migration and innovation: Does cultural diversity matter for regional R&D activity? *Papers in Regional Science*, 89(3):563–585.

Orford, S. (2002). Valuing locational externalities: A GIS and multilevel modelling approach. *Environment and Planning B: Planning and Design*, 29(1):105–127.

Ossokina, I. V. and Verweij, G. (2015). Urban traffic externalities: Quasi-experimental evidence from housing prices. *Regional Science and Urban Economics*, 55:1–13.

Ostendorf, W. J. M. and Musterd, S. (1996). Groeikernen en compacte steden. Nieuwe grenzen voor de ruimtelijke ordening. *Planologisch Nieuws*, 16(2):91–101.

Ottaviano, G. and Peri, G. (2006). The economic value of cultural diversity: evidence from US cities. *Journal of Economic Geography*, 6(1):9–44.

Ozgen, C., Nijkamp, P., and Poot, J. (2011). Immigration and Innovation in European Regions. *Tinbergen Institute Discussion Paper*, (112/3).

Pasidis, I., Viladecans-Marsal, E., and Garcia-López, M.-À. (2015). Express delivery to the suburbs. Transport Infrastructure and European cities. *ieb.ub.edu*.

Percoco, M. (2016). Highways, Local Economic Structure and Urban Development. *Journal of Economic Geography*, 16(5):1035–1054.

Pollakowski, H. O. and Wachter, S. M. (1990). The Effects of Land-Use Constraints on Housing Prices. *Land Economics*, 66(3):315–324.

Priemus, H. and Louw, E. (2002). Recovery of land costs: a land policy instrument missing in the Netherlands? *European Journal of Housing Policy*, 2(2):127–146.

Quigley, J. M. and Raphael, S. (2005). Regulation and the high cost of housing in California. *American Economic Review*, 95(2):323–328.

Raspe, O. and Van Oort, F. G. (2006). The Knowledge Economy and Urban Economic Growth. *European Planning Studies*, 14(9):1209–1234.

Rietveld, P. and Bruinsma, F. R. (1998). *Is transport infrastructure effective?: transport infrastructure and accessibility: impacts on the space economy*. Springer, Berlin Heidelberg.

Rietveld, P. and Wagtendonk, A. J. (2004). The location of new residential areas and the preservation of open space: experiences in the Netherlands. *Environment and Planning A*, 36(11):2047–2063.

Roback, J. (1982). Wages, Rents, and the Quality of Life. *The Journal of Political Economy*, 90(6):1257–1278.

Robinson, P. (1988). Root-N-consistent semiparametric regression. *Econometrica: Journal of the Econometric Society*, 56(4):931–954.

Rodríguez-Pose, A. and Ketterer, T. D. (2012). Do Local Amenities Affect the Appeal of Regions in Europe for Migrants?*. *Journal of Regional Science*, 52(4):535–561.

Rodríguez-Pose, A. and Vilalta-Bufí, M. (2005). Education, migration, and job satisfaction: The regional returns of human capital in the EU. *Journal of Economic Geography*, 5(5):545–566.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *The journal of political economy*, 82(1):34–55.

Saiz, A. (2010). The Geographic Determinants of Housing Supply. *The Quarterly Journal of Economics*, 125(3):1253–1296.

Saiz, A. and Wachter, S. (2011). Immigration and the Neighborhood. *American Economic Journal: Economic Policy*, 3(2):169–188.

Schwanen, T., Dieleman, F. M., and Dijst, M. (2001). Travel behaviour in Dutch monocentric and policentric urban systems. *Journal of Transport Geography*, 9(3):173–186.

Schwanen, T., Dieleman, F. M., and Dijst, M. (2004). Policies for urban form and their impact on travel: the Netherlands experience. *Urban Studies*, 41(3):579–603.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis.* Chapman and Hall, London.

Simon, C. J. (1998). Human Capital and Metropolitan Employment Growth. *Journal of Urban Economics*, 43(2):223–243.

Smaal, M. L. (2012). *Politieke strijd om de prijs van automobiliteit: De geschiedenis van een langdurend discours : 1895-2010.* Eburon, Delft.

Smersh, G. T. and Smith, M. T. (2000). Accessibility Changes and Urban House Price Appreciation: A Constrained Optimization Approach to Determining Distance Effects. *Journal of Housing Economics*, 9(3):187–196.

Spiekermann, K. and Neubauer, J. (2002). European accessibility and peripherality: Concepts, models and indicators. *Nordregio WP*.

Suedekum, J., Wolf, K., and Blien, U. (2014). Cultural Diversity and Local Labour Markets. *Regional Studies*, 48(1):173–191.

Theebe, M. (2004). Planes, trains, and automobiles: the impact of traffic noise on house prices. *The Journal of Real Estate Finance and Economics*, pages 209–234.

Tillema, T., Hamersma, M., Sussman, J. M., and Arts, J. (2012). Extending the Scope of Highway Planning: Accessibility, Negative Externalities and the Residential Context. *Transport Reviews*, 32(6):745–759.

Turner, M. A., Haughwout, A., and Van der Klaauw, W. (2014). Land Use Regulation and Welfare. *Econometrica*, 82(4):1341–1403.

Van Duijn, M., Möhlmann, J., Mulalic, I., and Rouwendal, J. (2014). *Models of household location and urban amenities.* HELP International report nr. 3-Higher Educated Location Preferences (HELP).

Van Duijn, M. and Rouwendal, J. (2013). Cultural heritage and the location choice of Dutch households in a residential sorting model. *Journal of*

Economic Geography, 13(3):473–500.

Van Ommeren, J. N. and Van der Vlist, A. J. (2016). Households' willingness to pay for public housing. *Journal of Urban Economics*, 92:91–105.

Van Oort, F. G. (2004). *Urban growth and innovation: Spatially bounded externalities in the Netherlands*. Gower Publishing, Ltd.

Veenendaal, A. J. (1996). Railways and the Coming of Road Transport in the Netherlands, 1919-1940. *Business and Economic History*, pages 293–300.

Vermeulen, W. and Rouwendal, J. (2007). Housing supply in the Netherlands.

Vermeulen, W. and Rouwendal, J. (2014). on the Value of Foregone Open Space in Sprawling Cities. *Journal of Regional Science*, 54(1):61–69.

Vickerman, R., Spiekermann, K., and Wegener, M. (1999). Accessibility and economic development in Europe. *Regional Studies*, 33(1):1–15.

VROM (1993). Vierde nota over de ruimtelijke ordening extra, deel 4: Planologische kernbeslissing nationaal ruimtelijk beleid. Technical report, Ministerie van Volkshuisvesting, Ruimtelijke ordening en Milieu, SdU, Den Haag.

VROM (2004). Nota ruimte. Ruimte voor ontwikkeling. Technical report, Ministerie van Ministerie van Volkshuisvesting, Ruim- telijke ordening en Milieu, Ministerie van Landbouw, Natuurbeheer en Visserij, Ministerie van Verkeer en Waterstaat and Ministerie van Economische Zaken, SdU, Den Haag.

Wang, F. T. and Zorn, P. M. (1997). Estimating house price growth with repeat sales data: what's the aim of the game? *Journal of Housing Economics*, 6(2):93–118.

Wang, Z., De Graaff, T., and Nijkamp, P. (2016). Cultural Diversity and Cultural Distance as Choice Determinants of Migration Destination. *Spatial Economic Analysis*, 11(2):176–200.

Wilhelmsson, M. (2000). The Impact of Traffic Noise on the Values of Single-family Houses. *Journal of Environmental Planning and Management*, 43(6):799–815.

Wooldridge, J. M. (2002). *Econometric Analysis of cross section and panel data*. MIT Press Books. MIT Press.

Yatchew, A. (1998). Nonparametric regression techniques in economics. *Journal of Economic Literature*, 36(2):669–721.

Yiu, C. Y. and Wong, S. K. (2005). The effects of expected transport improvements on housing prices. *Urban Studies*, 42(1):113–125.

Summary

Urban spatial policies guide the development of cities and regions. These policies aim to improve and optimize economic and social welfare, for instance, by permitting development and designating protected areas, improving transportation accessibility and ensuring the provision of essential or attractive amenities in a region. However, interventions in the spatial setting may often result in ineffective, or sometimes unexpected economic outcomes. It may also be that the implementation of a certain policy is effective in achieving its goal, but lead to unintended indirect economic consequences.

In this dissertation I explore how the implementation of various urban spatial policies impacted local economic conditions. In particular, I focus on the effects of improved road accessibility, spatial planning restrictions and development of urban amenities, on local housing and land markets, distribution of population and suburbanization, and the location decisions of individuals. To address these questions I make use of several advanced econometric methods to analyze unique databases on housing and land transactions, historic spatial data on transportation networks and land use, and micro-level data on individuals' revealed preferences for residential locations and their demographic characteristics.

The first section of the dissertation is dedicated to the evaluation of the effects of accessibility and spatial planning policies. The first question I examine is how the development of new highways affects local house prices. Examination of a large number of housing transactions from the Netherlands shows that improved road accessibility through new highways is valued positively, except in areas immediately adjacent to the new roads, and that it is reflected in an increase in local residential real estate values. The findings also provide evidence of an anticipation effect, as real estate values already experience an increase before the construction of the highways is completed.

A substantial body of literature has also established that highway accessibility drives suburbanization and outward expansion of cities. In the third chapter we examine how this prediction is changed when strict land development regulation exists, as was during the expansion of the Dutch highway

network in the 1960's. The findings indicate that when land development is restricted in the surroundings of cities, new highways divert population growth to locations further away from central cities, resulting in a large scale sprawl which 'leapfrogged' over the restricted zones to peripheral towns.

Land development regulation and land use regimes are common forms of planning policies which ensure the orderly development of urban areas. However, when different land uses are subject to different regimes, and when conversion between designated land uses is almost perfectly restricted, the consequence may be a segmentation of the land market based on land use designation. In the fourth chapter we examine how strict development restrictions and differences in policy regimes in the Netherlands between residential and commercial land uses result in divergence between the land markets, which is reflected in a substantial divergence in values of undeveloped lands.

The findings of this section emphasize that while improved accessibility is often expected to increase real estate values and population levels, exceptions may apply under certain spatial or policy conditions. Moreover, strictly enforced planning policies may intervene with free market outcomes of improved accessibility and land markets. Understanding how the presence of strict spatial planning regulations interferes with expected suburbanization and urban expansion processes can support better-informed policy decisions. Similarly, understanding that certain land regulations may result in land market segmentation and divergence in land values can support policy decisions which are able to consider these indirect outcomes and to assess whether they are desirable or socially optimal.

The second section of the dissertation is dedicated to the effects of urban policies which focus on the provision of urban amenities. Urban amenities are desirable or useful features of a city or a neighborhood, which are valued by individuals and affect their location decisions. In order to increase the attractiveness of an area, and its local real estate values, public resources are often allocated to develop and maintain amenities which are perceived as positive. In the Netherlands, the relative abundance of water is often utilized for this purpose. The fifth chapter includes an investigation of the value of proximity to water, as reflected in transaction prices of model homes. We show that the effect of proximity to water on housing value is lower and more local than what was previously found.

The value which households attach to local amenities is not identical, but closely depends on socioeconomic and demographic characteristics. Differences in willingness-to-pay can therefore influence households' residential decisions and neighborhoods composition. In chapter six I follow a sorting model framework to examine how urban features, such as cultural heritage, open space and the presence of migrants, are valued by individuals who differ in characteristics, particularly origin and higher-education attainment. The findings show that highly-educated migrants attach a higher value for areas richer in amenities, and that the presence of migrants is valued positively by all groups, but this value decreases as the share of migrants increases.

The findings of this section shed light on the value which households attach to different urban amenities, and how the provision of certain amenities influence their residential location decisions. This can be used to better guide investments in development and maintenance of urban amenities, and avoid inefficient allocation of public resources. Moreover, differences between heterogeneous groups in willingness-to-pay values also reveal their preferences for certain amenities, and are therefore essential in understanding sorting processes of households in an urban area.